63-3-4

ASD-TDR-63-59

DEVELOPMENT OF HIGH TEMPERATURE SEMICONDUCTOR DEVICES

TECHNICAL DOCUMENTARY REPORT NO. ASD-TDR-63-59

February 1963

Electronic Technology Laboratory
Aeronautical Systems Division
Air Force Systems Command
Wright-Patterson Air Force Base, Ohio

Project 4460, Task No. 446002

(Prepared under Contract No. AF33(600)-37267
by Radio Corporation of America,
Semiconductor and Materials Division,
Somerville, New Jersey
Authors: H. Becke, W. A. Bosenberg, N. Ditrick, F. L. Vogel)

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Qualified requesters may obtain copies of this report from the Armed Services Technical Information Agency, (ASTIA), Arlington Hall Station, Arlington 12, Virginia.

This report has been released to the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C., for sale to the general public.

Copies of this report should not be returned to the Aeronautical Systems Division unless return is required by security considerations, contractual obligations, or notice on a specific document.

27-4068, 200, 5-8-43

FOREWORD

This report was prepared by the Radio Corporation of America, Semiconductor and Materials Division, Somerville, New Jersey on Air Force Contract No. AF33(600)-37267 under Task No. 446002 of Project No. 4460 "Development of High Temperature Semiconductor Devices". The work was administered under the direction of the Electronic Technology Laboratory, Aeronautical Systems Division. Mrs. E. Tarrants was task engineer for the Laboratory.

The studies presented in this report began in May 1958 and were concluded in July 1962.

This report concludes the work on Contract No. AF33(600)-37267.

ABSTRACT

Gallium arsenide was chosen as the semiconductor material best suited to meet the needs of high-temperature operation because of its wide band gap and high electron mobility. Even though gallium arsenide has the best potential for high-temperature operation, it does have a few limitations. One of these being that there are a number of major differences between elemental and compound semiconductors. These differences include such things as, stochiometry, positive and negative crystal faces, and decomposition of the semiconductor material during processing. All of these are unique to the compound semiconductor and require special attention. One major limitation of gallium arsenide is band to band carrier recombination which results in very short minority carrier lifetimes.

A number of different devices were developed under this contract. These include: switching diodes, low-power rectifiers, medium-power rectifiers, high-power rectifiers, zener diodes, tunnel diodes, unipolar transistors, switching transistors, power transistors and solid ceramic circuits. Of these, the most successful were tunnel diodes and varactor diodes, since they are able to fulfill the requirements of some specific applications not filled by similar germanium and silicon devices.

Switching diodes with a recovery time of one nanosecond and an operating temperature limit of 400°C were made. These are better than corresponding silicon diodes, especially at elevated temperatures. Low-, medium - and high-power rectifiers with an operating temperature of 400°C, which is much higher than can be achieved with similar silicon devices, were fabricated. Unipolar transistors were investigated, but because of the relatively crude device technology in this area, full advantage was not taken of gallium arsenide's potential. Switching and power transistor were also made with satisfactory characteristics (except current gains which were low). The high frequency characteristics of these transistors, however, are good.

PUBLICATION REVIEW

Publication of this technical documentary report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

FOR THE COMMANDER

RICHARD D. ALBERTS
Chief, Molecular Electronics
Branch
Electronic Technology Laboratory

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

LIST OF FIGURES

LIST OF FIGURES (Cont.)

LIST OF FIGURES (Cont.)

LIST OF TABLES

I. TECHNICAL DISCUSSION

A. Introduction-Gallium Arsenide as a Semiconductor Material

The present trend in semiconductor technology is directed toward the development of devices capable of operating at higher and higher frequencies. In semiconductor devices this calls for smaller sizes, higher current densities and, hence, higher power dissipation capability. Furthermore; military and space applications necessitate packing large numbers of components into the smallest volumes possible and also operation in ambients which are often heated aerodynamically or by proximity of the power source or simply by the high density of electronic components. As a result of these conditions, it is becoming necessary for the devices to withstand unusually high operating temperatures. Since germanium and silicon are presently approaching the limits of their capabilities, they cannot be relied upon for future improvements. At least one compound semiconductor, gallium arsenide, shows promise of not only combining the favorable properties of germanium and silicon but of exceeding the best combined features of both. Figure 1 shows the maximum operating temperature, the band gap, and the electron mobility for several semiconductor materials. Both gallium arsenide and indium phosphide have high-temperature capability along with high mobilities (which is related to high-frequency performance). Other semiconductors, such as gallium phosphide, silicon carbide, and carbon are capable of going to even higher temperatures, but unfortunately have very low mobilities. As a result, both gallium arsenide and indium phosphide devices were investigated at the start of this contract.

A comparison of significant electrical properties (see Table I) reveals that while gallium arsenide and indium phosphide are similar, the electrical values are uniformly in favor of gallium arsenide. From the aspect of

Manuscript released by the author January, 1963 for publication as an ASD Technical Documentary Report

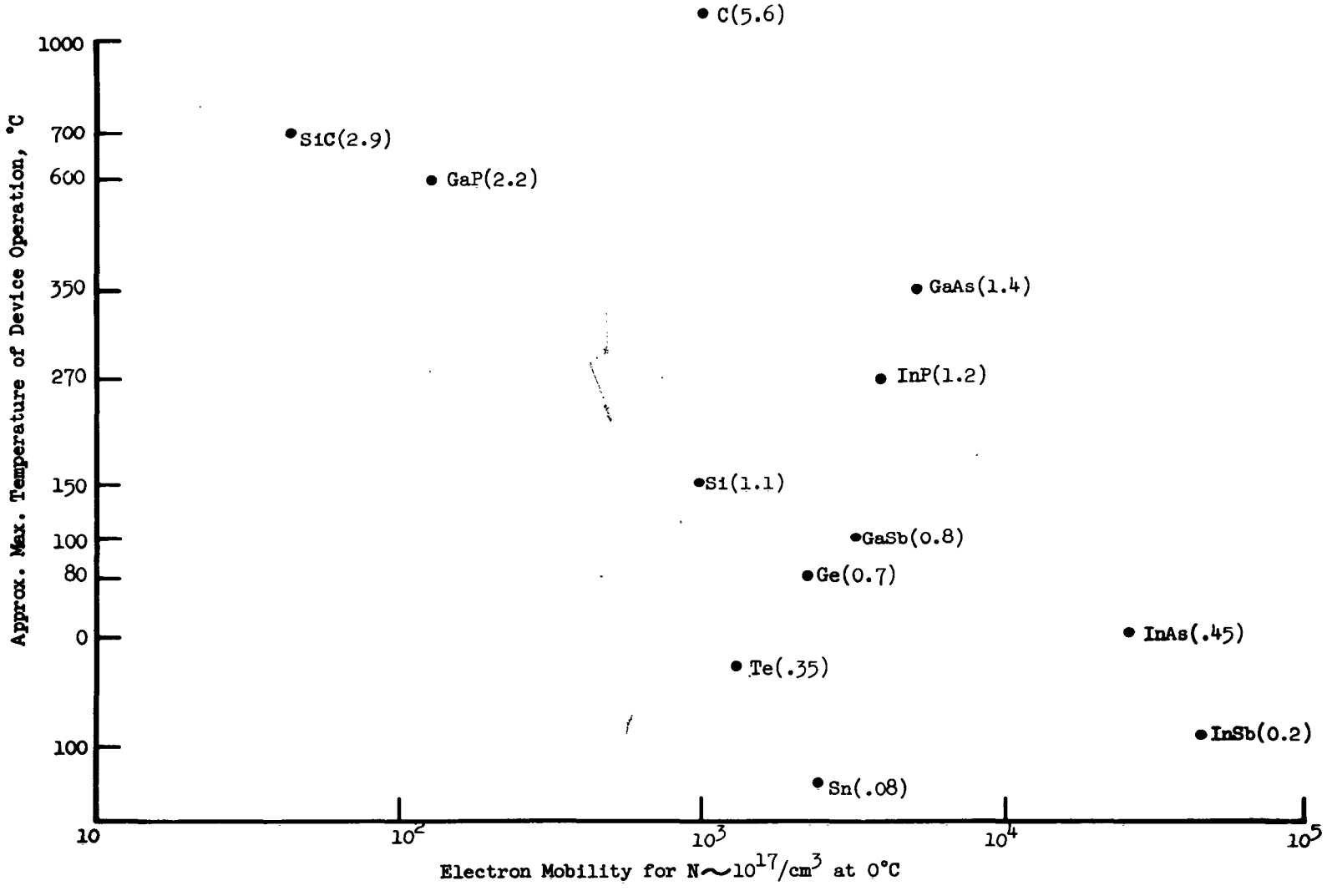


FIGURE 1 MAXIMUM OPERATING TEMPERATURE, BAND GAP AND ELECTRON MOBILITY OF SEVERAL SEMICONDUCTOR MATERIALS

crystal growing the lower melting point of indium phosphide would make growth easier by minimizing crucible reactions. On the other hand, the high phosphorus pressure, at the indium phosphide melting point, could lead to disastrous difficulties. Consideration of all of these points in conjunction with the more advanced state of gallium arsenide technology led to the decision that effort should be concentrated on gallium arsenide.

TABLE I

COMPARISON OF SIGNIFICANT ELECTRICAL PROPERTIES IN GALLIUM ARSENIDE AND INDIUM PHOSPHIDE

Material	Eg (300°K), ev	Me (300°K), cm^2/V sec	Mh (300°K), cm^2/V sec	Melting point, °C	Partial pressure at melting point, Atm.
GaAs	1.4	8500	420	1237	P_{AS} =.9
InP	1.29	4600	150	1062	P_P =60

Experience with germanium and silicon has shown that there are four semiconductor properties (shown in Table II) which are most important in determining the significant device characteristics.

TABLE II

DESCRIPTIVE TERMS, SYMBOLS AND DEFINITIONS OF SEMICONDUCTOR PROPERTIES

Descriptive Term	Symbol	Definition
Band Gap	E_g	The energy required to produce an electron-hole pair.
Impurity Activation Energy	E_1	The energy required to produce an electron or a hole.
Electron and Hole Mobility	μ_n, μ_p	The drift velocity of an electron or a hole in an electric field of unity.
Dielectric Constant	ϵ	(Conventional)

Table III shows the important device-material relationship which formed the basis for the selection of gallium arsenide for high performance device applications.

TABLE III

SEMICONDUCTOR PROPERTIES AND DEVICE PERFORMANCE

	Category	Semiconductor Property	Approximate Relation
Operating Temperature Range	Upper Temperature Limit T_u	Band Gap E_g	$T_u \alpha E_g$
	Lower Temperature Limit T_l	Impurity Activation	$T_l \alpha E_l$
Upper Frequency Limit F (Switching Speed)	Unipolar Devices (unipolar transistors; high frequency and switching diodes	The higher of the two mobilities μ high; Dielectric constant	$F \alpha \frac{\mu \text{ high}}{\epsilon^{1/2}}$
	Bipolar Devices (Transistors, n-p-n-p switches)	Electron and Hole Mobility, μ_n and μ_p; Dielectrin constant	$F \sim \frac{\mu_n \times \mu_p}{\epsilon^{1/2}}$

A high band gap and high mobilities in conjunction with low impurity activation energy and a low dielectric constant are generally desired for high-performance semiconductor devices. High band gaps permit operation at higher temperatures. Silicon, for instance, has a band gap of 1.1ev and, therefore, is limited to temperatures, below 200°C, whereas gallium arsenide's band gap of 1.4ev would allow device operation up to 400°C. On the other hand, it is desirable, almost necessary, that a high-temperature device operate at room temperature as well. This behavior, controlled by the impurity activation energy, is satisfactory in gallium arsenide where the impurity activation energies are less than 0.1ev, and the impurities are easily activated at room temperature. Gallium arsenide is in sharp contrast with silicon carbide, which is capable of making devices for operation at

very high temperatures, but these devices will not function at room temperature. A comparison of the frequency characteristics of germanium, silicon and gallium arsenide bipolar and unipolar devices is given in Figure 2 and Figure 3. Figure 2 refers to unipolar devices. The term unipolar refers to the fact that carriers of only one type participate in the active electronic mechanism of the device. The relative frequency values are based on the assumption that a germanium unipolar device will operate at a frequency of 1 kmc at room temperature. Figure 3 shows the equivalent plot for bipolar transistors. In this type of device both charge carriers participate.

A description of some of the other significant properties of gallium arsenide will aid in clarifying the evolution and progress of device fabrication as described in later sections. Gallium arsenide is a III-V compound semiconductor with a zinc-blende crystal structure identical to that of germanium and silicon. The only difference is that alternate sites are occupied by gallium and arsenic atoms (see Figure 4). Some marked differences in crystal properties stem from this duel occupancy in the lattice structure and are reflected in device technology. It is clear from the figure that the opposite (111) planes differ, one consisting exclusively of gallium atoms (the + (111) plane) and the other of arsenic atoms (the $(\bar{1}\bar{1}\bar{1})$ plane). When these two planes are etched, the arsenic face takes on a high polish while deep etch pits appear on the gallium face. Device technology has evolved by recognizing these differences and providing a means of establishing wafer orientation early in device processing. Polishing etches have proven effective for this purpose.

The energy band structure for gallium arsenide, shown in Figure 5, gives some of the important properties of the material. The large energy

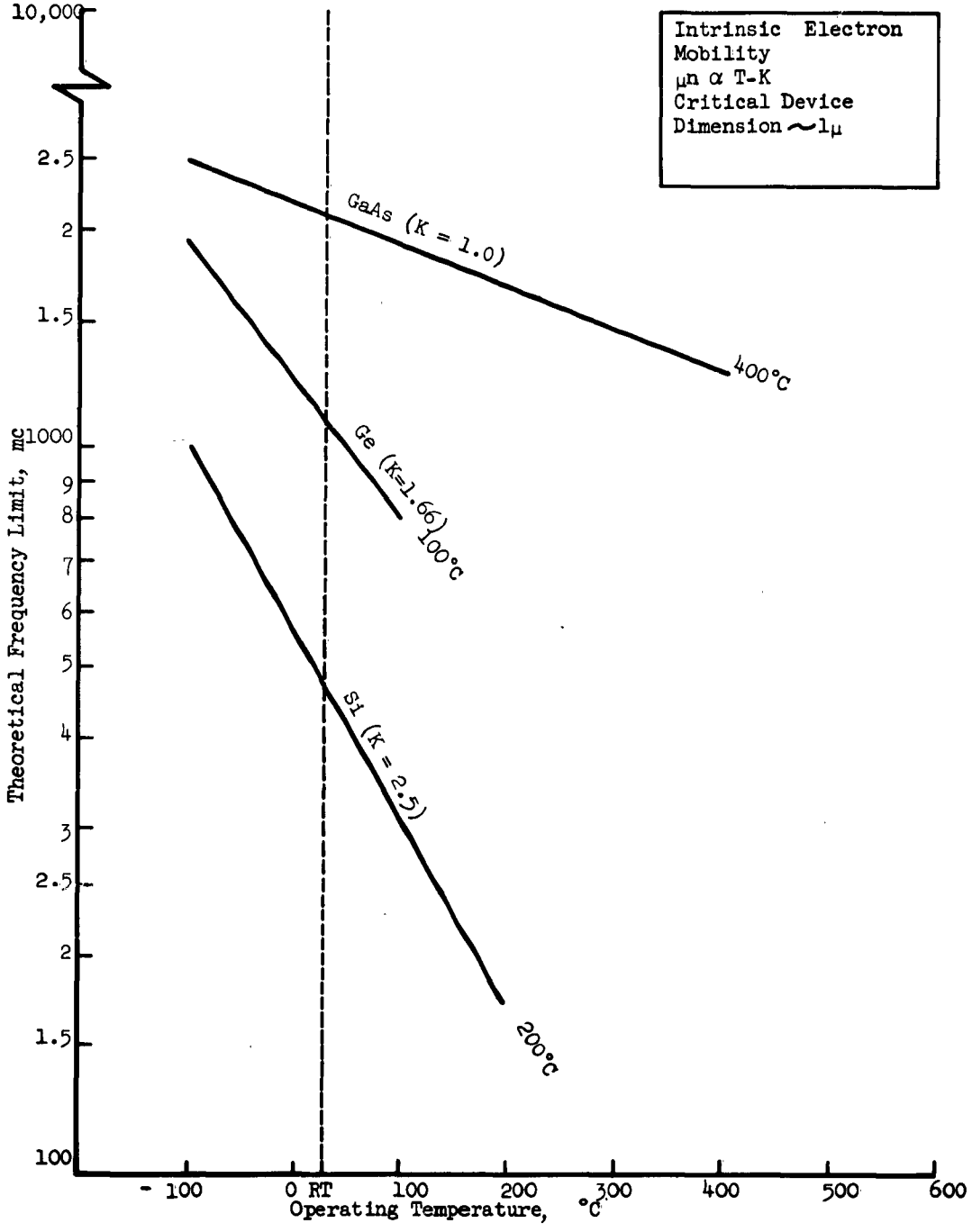


FIGURE 2 RELATIVE FREQUENCY LIMIT VS. TEMPERATURE OF UNIPOLAR DEVICES IN GALLIUM ARSENIDE, SILICON AND GERMANIUM SEMICONDUCTORS

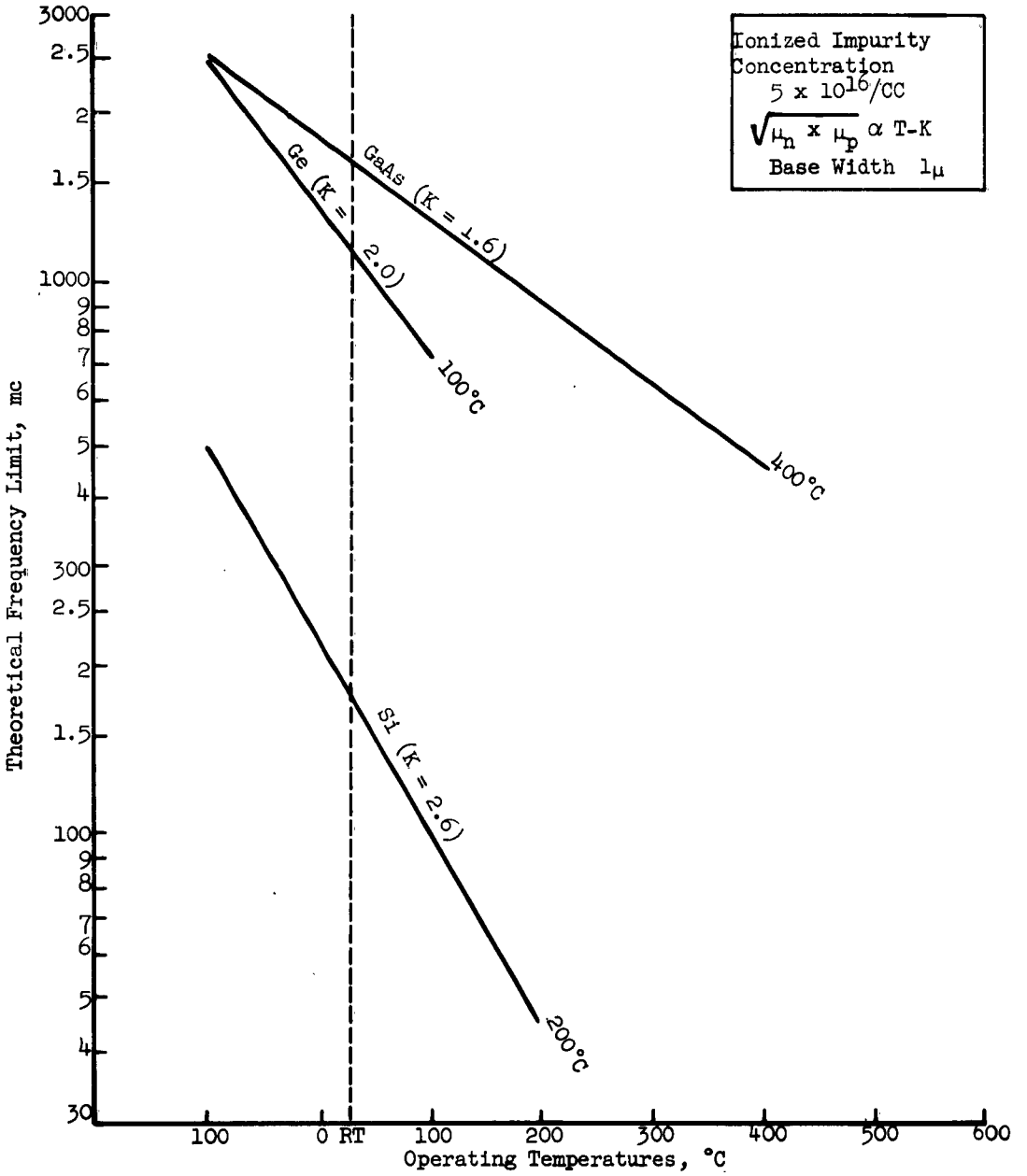


FIGURE 3 RELATIVE FREQUENCY LIMIT VS. TEMPERATURE OF BIPOLAR DEVICES IN GALLIUM ARSENIDE, SILICON AND GERMANIUM SEMICONDUCTORS

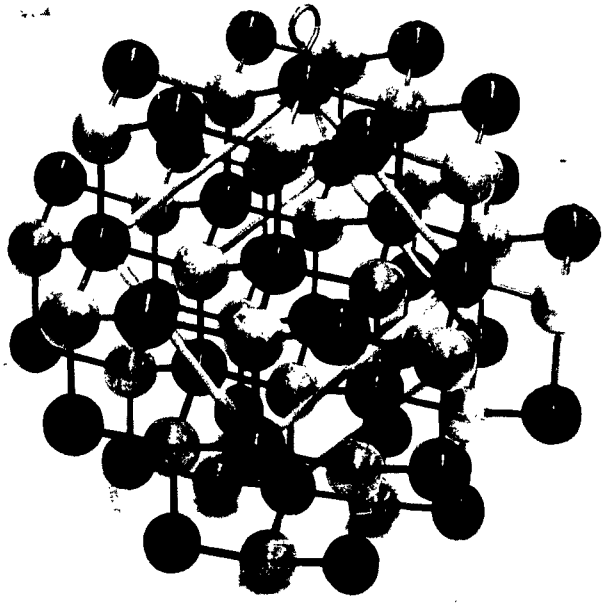


FIGURE 4 CRYSTAL STRUCTURE OF GALLIUM ARSENIDE
(Black Balls - Arsenic White Balls - Gallium)

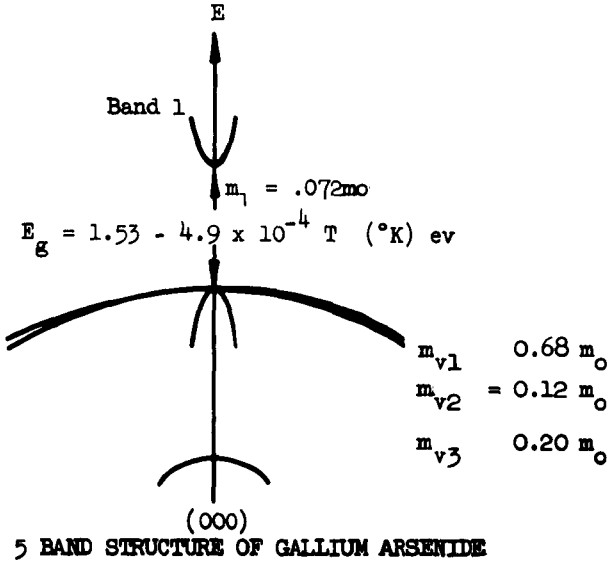


FIGURE 5 BAND STRUCTURE OF GALLIUM ARSENIDE

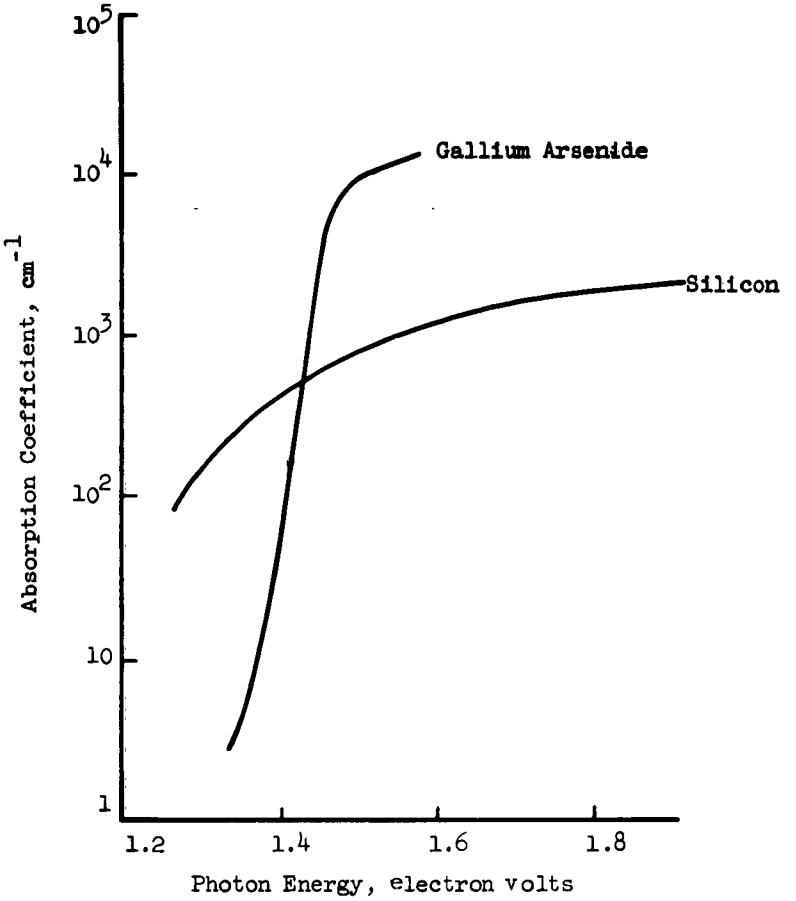


FIGURE 6 OPTICAL ABSORPTION IN SILICON AND GALLIUM ARSENIDE CRYSTALS

cannot undergo severe heat treatments as shown in the next paragraph.

A property of available gallium arsenide crystal which plays a significant role in device capability is "thermal conversion". When lightly doped gallium arsenide crystal is heat treated, during diffusion for example, the net carrier concentration drops abruptly, the mobility decreases and resistivity values of 10^4 to 10^6 ohm-cm result. It was found possible to use radiant heating, copper leaching, and other techniques to limit the extent of conversion but, in a practical sense, this thermal conversion property

gap

$$E_g = 1.53 - 4.9 \times 10^{-4} T\ (^\circ K)\ ev$$

separates the electron population with a very low effective mass (0.072mo) and a hole population in a multiple degenerate band, with a higher, but still small, effective mass. The conduction band minimum and the valence band maximum occur at the same point in momentum space. In other words, an electron-hole pair can recombine directly across the energy gap without requiring momentum transfer from the lattice. This is very different from the situation in germanium and silicon where recombination through trapping centers predominates. This has important implications for the lifetime of minority carriers in gallium arsenide crystals as will be seen below.

Some basic information on carrier lifetime can be derived from consideration of the optical absorption curves for gallium arsenide and silicon which are shown in Figure 6. The abruptness of the absorption edge in gallium arsenide is indeed striking. The shape of the curve implies that for an energy only slightly in excess of the bandgap value, a hole-electron pair is created easily. From detailed valence arguments, one can show that the converse process, direct recombination of a hole-electron pair, also occurs easily. This was recognized and the shape of the absorption curve was inserted into a relation by von Rossbroeck and Shockley(ref.1) to calculate the recombination rate for minority carriers. The result of this calculation sets a maximum value for the minority carrier lifetime as shown in Figure 7. The lifetime is inversely proportional to the majority carrier density and, at the 10^{17} carriers cm^3 doping level, for example, the lifetime for direct recombination is 2×10^{-8} seconds. Measured lifetimes are generally at least one order of magnitude smaller in available crystals. Thus, only limited improvements in the minority carrier lifetime of gallium arsenide are possible. Devices requiring longer lifetimes must use very low impurity densities, and, therefore,

ref. 1 Von Roosbroeck and Shockley, "Photon Radiative Recombination of Electrons and Holes in Germanium", Phys. Rev. 94, p. 1558, 1954

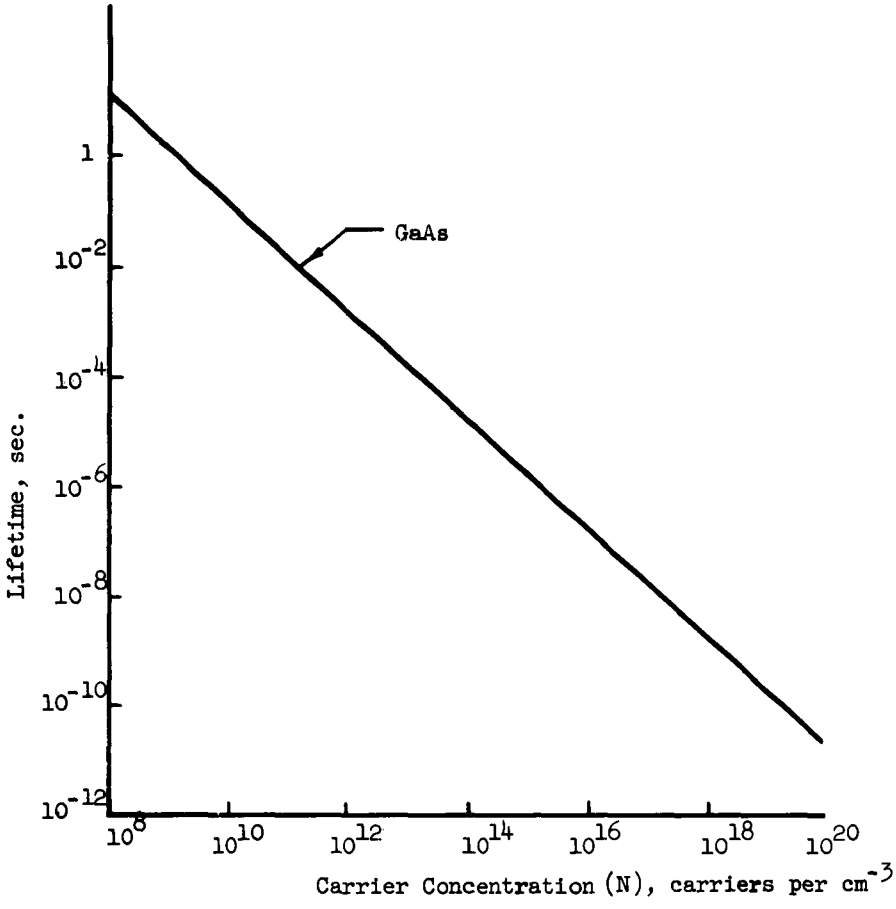


FIGURE 7 LIFETIME FOR DIRECT RECOMBINATION GALLIUM ARSENIDE

limits the minimum doping level of gallium arsenide crystal to 1 to 3×10^{16} carriers/cm^3. The effect of this limit will be discussed in the various device sections.

Gallium arsenide, then, is characterized by a wide band-gap, a high electron mobility, and a short minority carrier lifetime. Some of the other material properties of gallium arsenide are shown together with the values for germanium and silicon in Table IV for purposes of comparison.

From the foregoing analysis it is clear that gallium arsenide is the most promising material in terms of high-frequency, high-temperature performance.

TABLE IV

PROPERTIES OF GERMANIUM, SILICON AND GALLIUM ARSENIDE

Properties	Ge	Si	GaAs
Melting Point (°C)	936	1420	1238
Density (g/cm^3)	5.3232	2.328	5.312
Liquid Density (g/cm^3 at M.P.)	5.51		5.45
Lattice Constant (Å)	5.6576	5.431	5.654
Dist. Betw. Nearest Neighbors (Å)	2.45	2.35	2.45
Atoms per cm^3 ($\times 10^{22}$)	4.42	4.99	4.43
Thermal Coeff. of Expansion (/°C)	6.1×10^{-6}	4.2×10^{-6}	5.93×10^{-6}
Pressure at M.P. (atm)	10^{-9}	10^{-6}	0.9
Thermal Conductivity (watt units)	0.63	.84	0.37
Specific Heat (cal/gm)	.074	.181	0.086
Latent Heat of Fusion (kcal/mole)	8.3	9.45	21 ± 3
Dielectric Constant	15.7	11.7	11.1
Elastic Moduli ($\times10^{12}$ dynes/cm^2) C_{11}	1.298	1.674	1.188
C_{12}	.488	.652	0.538
C_{44}	.673	.796	.594
Volume Compressibility ($\times10^{-12}$ cm^2/degree)	1.3	.98	
Magnetic Susceptibility (cgs)	$-.12\times10^{-6}$	$-.13\times10^{-6}$	
Debye Temperature (°K)	362	652	
Emissivity (1200°C)	.5	.5	
Band Gap (25°C)	.67	1.106	1.40
Temp. Dep. of Band Gap (ev/°C)	-4.1×10^{-4}	-4.4×10^{-4}	-4.9×10^{-4}
Electron Lattice Mobility (cm^2/v-sec)	3950	1900	12,000
Hole Lattice Mobility (cm^2/v-sec)	1900	425	450
Temp. Dep. Electron Lattice Mobility	$T^{-1.66}$	$T^{-2.5}$	T^{-1}
Temp. Dep. Hole Lattice Mobility	$T^{-2.33}$	$T^{-2.7}$	$T^{-2.1}$
Electron Effective Mass Ratio	0.55	1.1	.072
Hole Effective Mass Ratio	0.37	0.59	0.5
Intrinsic Electrons (25°C) (cm^{-3})	2.4×10^{13}	1.5×10^{10}	1.4×10^{16}
Intrinsic Resistivity (25°C) (ohm-cm)	46	2.3×10^{5}	3.7×10^{8}

Several other advantages which may contribute significantly to the importance of gallium arsenide as a semiconductor device material, were not fully appreciated at the early stages of this project. For example, gallium arsenide ranks very high as a material to be used for the conversion of solar energy by a photovoltic device, and as a practical matter, will probably be the material which gives the highest reliable conversion efficiency. Likewise, since the quantum efficiency of forward current in gallium arsenide p-n junctions is high, close to 100%, this material offers the possibility of making a truly unique laser which would operate by the injection of minority carriers rather than by optical pumping.

B. Growth of Gallium Arsenide Crystals

During the initial stages of the contract a number of techniques, for the growth of high-purity, single-crystal gallium arsenide were investigated. These techniques included

1) Horizontal Bridgman
2) Magnetically Coupled Czochralski
3) Vertical Bridgman
4) Excess gallium solution
5) Excess arsenic solution
6) Vapor deposition

The last three methods were investigated in an attempt to obtain material with a greater purity than was possible with melt techniques. This higher purity was considered possible since the presence of a hot crucible was considered the major source of contaminating impurities. The resulting material showed no significant or reproducible reduction in impurities or improvement in electrical properties and was also found poor from the viewpoint of size and crystallinity. Further efforts in these directions were not considered sufficiently promising for continued investigation.

The vertical Bridgman method was selected as the second source of gallium arsenide crystal, since it promised to be a quick and relatively simple source of obtaining single crystal material. Unfortunately, the crystals were quite strained due to the constraints imposed upon them during freezing by the quartz tube. No purity or property improvement over the horizontal Bridgman technique was apparent.

The greatest success in producing single crystals on a large scale was obtained in the horizontal Bridgman furnaces. In this method the strain incurred by growth in a boat was controllable, and after growth techniques were refined all of the gallium and arsenic was converted to single crystal.

Furthermore, the equipment necessary is comparatively inexpensive so that large scale production could easily be attained.

The impurities in the grown crystals were usually in the neighborhood of a few parts per million, which was considered suitable for device fabrication, but the mobilities seldom approached the theoretical values.

The advantages of this method were, therefore, 1) high yield of good crystal, 2) low contamination level, and 3) simplicity and low cost. The disadvantages were 1) lower than theoretical mobility and 2) high dislocation density.

The two disadvantages were attributed to boat constraint and boat wetting and attack by the gallium and gallium arsenide. These difficulties could be relieved by pulling the crystal from a melt. Thus, there would be no boat constraint. In addition, a crucible that was impervious to attack could also be used.

A magnetically coupled Czochralski furnace was designed and used for the production of low dislocation material. Only recently were sufficiently pure crucibles available for growth of high mobility crystals. This process is, however, more complex and requires more highly skilled operators than the horizontal Bridgman technique.

At present both the horizontal Bridgman and magnetically coupled Czochnalski techniques are developed to the point of being useful in production.

C. Switching Diodes and Low-Power Rectifiers

1. Design Considerations

The basic electrical parameters for diodes and rectifiers include the reverse breakdown voltage, V_B, the leakage current at high temperature, I_R, and the average rectified current, I_{AV}. For switching diodes, the reverse recovery time is also important. These parameters are determined by the device design parameters which include the series resistance of the diode, R_S, the thermal resistance between the diode junction and the ambient and the junction capacitance, C. The design of the device involves all of those parameters.

a. Breakdown Voltage

The breakdown voltage of a p-n junction (for $V_B > 10$ volts) is generally determined by avalanche breakdown in which the passage of minority carriers through the high field depletion layer results in an accumulated minority carrier energy sufficient to ionize a hole-electron pair. These carriers, in turn, also traverse the depletion layer giving rise to further pairs until, at V_B, the process runs away. For a given voltage, the amount of doping controls the electric field in the depletion layer. Since avalanche breakdown is dependent upon the electric field, lighter doping produces higher breakdown voltage. The thermal conversion problem is the factor which limits the minimum doping level in gallium arsenide. Because of this problem, material with a doping level of 1 to 3 x 10^{16} carriers/cm^3 is required. The relationship between breakdown voltage and doping for nearly abrupt junctions in n-type gallium arsenide is shown in Figure 8 (ref.2). A similar

ref. 2. H.Kressel, A. Blicher and L.H. Gibbons, Jr. to be published in Proc. IRE

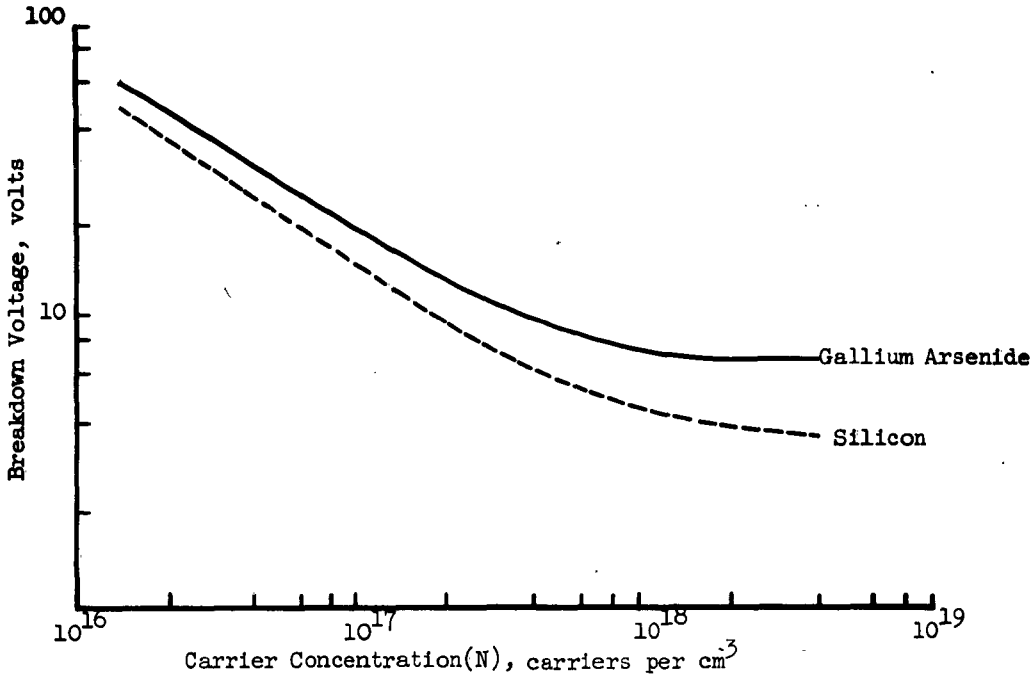


FIGURE 8 BREAKDOWN VOLTAGE OF GALLIUM ARSENIDE AND SILICON p-n JUNCTIONS AS A FUNCTION OF CARRIER CONCENTRATION

curve, for alloyed silicon diodes, has been plotted on the same graph. From this comparison it is evident that the breakdown electric field in gallium arsenide is larger than in silicon and that breakdown voltages of only about 50 volts can be achieved with gallium arsenide crystal that does not convert to p-type material.

It is a well known fact that diffused junctions exhibit larger breakdown voltages than abrupt junctions. As shown for silicon (ref.3) the breakdown voltage, V_B, is related to the concentration gradient, A, at the junction and the max. electric field, E_{max} by:

$$V_B = \text{const} \sqrt{E^3_{max}/A} \qquad (1)$$

ref.3 H.S. Veloric, M.B. Prince and M.J. Edger, "Avalanche Breakdown Voltage in Silicon Diffused P-N Junctions as a Function of Impurity Gradiant", J. Appl. Physics Vol. 27 pp 895-899 (1956)

By deeper diffusion, the concentration gradient, A, is reduced ($\sim 1/x_j$) and V_B increased. If E_{max} is assumed to be constant with doping level then V_B is related to the junction depth, x_j, by:

$$V_B \sim \sqrt{x_j} \qquad (2)$$

By diffusing, one takes advantage of the fact that the voltage across a junction is proportional to the movement of the ionized charge while the field is proportional to the total ionized charge. Any technique which moves the doping to the outer edge of the depletion layer gives an improved breakdown voltage.

Using this technique breakdown voltages of up to 90 volts have been obtained with n-type gallium arsenide having a carrier concentration of about 3×10^{16} carriers/cm^3.

b. pn Junction Characteristics

Germanium devices obey the following voltage-current characteristic:

$$I = I_o (\exp \alpha V - 1) \qquad (3)$$

which means that for $|V| \longrightarrow 1/\alpha$, $I = I_o$, a saturation behavior is obtained. In silicon devices Sah, Noyce and Shockley (ref.4) have used space charge generation and recombination to explain the non-saturation behavior of the V-I characteristics.

Similarly in gallium arsenide devices, much higher currents are observed in the reverse direction than can be attributed to the simple pn junction theory. Qualitatively the observed behavior can be explained by the Sah-Noyce-Shockley theory. Most gallium arsenide rectifiers produced were dominated by surface leakage at room temperature. However, a few were measured showing reverse currents as low as 10^{-11} amperes at 5-volts reverse bias. Since the junction

ref. 4 C.T. Sah, R.N. Nayee, and W. Shocklay, "Carrier Generation and Recombination in P-N Junctions and P-N Junction Characteristics", Proc. IRE, Vol. 45, pp. 1228-1243, 1957

area is about $3 \times 10^{-3} cm^{-3}$, the current density is 0.3×10^{-8} amp/cm^2. The Sah, Noyce and Shockley theory predicts a value of

$$J_{r_1 g} = \frac{q\, n_i\, w}{2\tau} \qquad (4)$$

where q = electron charge = 1.6×10^{-19} coulombs

n_i = intrinsic density in GaAs = 1.4×10^{6} carriers/cm^3

w = depletion layer width $\cong 5 \times 10^{-5}_{16}$ for a reverse voltage of 5 volts and a doping of 3×10^{16} carriers/cm^3.

τ = minority carrier lifetime $\sim 10^{-9}$ seconds

so that $J_{r_1 g} \cong 10^{-8}$ amperes/cm^2

which is in good agreement with the observed values.

In the forward direction all gallium arsenide devices show a relatively large amount of temperature independent excess current. This current, flowing to centers in the gap, masks the injection current. For the purposes of rectifier design, however, the forward bias region of greatest importance is dominated by the diode series resistance. Little of the conductivity modulation, so common in germanium and silicon devices, could be seen in gallium arsenide. This is due to the low lifetime for injected carriers. The series resistance for the typical 250-ma rectifiers is between one and two ohms. The bulk resistivity makes up about half of this, the contact resistance the other half.

c. Storage Time

The recovery time of a switching diode from forward injection, is described by Lax and Neustadter (ref.5) by:

$$\operatorname{erfc}\left(t_r / \tau\right) = \frac{1}{1 + I_r / I_f}$$

where: t_r = duration of the storage time

ref. 5 B.Lax and S.F. Neustadter "Transient Response of a P-N Junction" J. Appl. Physics, 25 P.1148 1954

τ = lifetime

I_f = initial forward current

I_r = constant reverse current

This relation assumes that the storage time for minority carriers will be larger than the R-C time constant (used for charging the junction capacitance through the series resistance of the diode). Switching times as short as 1 nanosecond were observed, but it was clear that the characteristic was dominated by charging time rather than by the decay of minority carriers. It is well known that doping with heavy metals, such as gold, can decrease the lifetime even more if need be. Gallium arsenide, then, unlike silicon or germanium, shows no minority carrier storage effects down to the nanosecond region.

d. Thermal Dissipation

Gallium arsenide diodes have a higher forward voltage drop than either germanium or silicon. This is mainly due to the larger band gap of gallium arsenide (1.4ev) as compared with germanium (0.67) and silicon (1.1ev). The higher the band gap the higher the possible operating temperature.

In addition, there is no appreciable forward conductivity modulation in gallium arsenide. At 250-ma rectified current, about 0.5 watts have to be dissipated, mostly as loss in the forward direction. As a result, the thermal resistance has to be made small enough so that the device can withstand this thermal dissipation. Values of about 50°C/watt are typical which indicate a safe limit of 25°C temperature rise.

This subject will be discussed further in the power rectifier section.

e. Package

A special hermetic package capable of withstanding elevated temperatures, is required for gallium arsenide switching diodes and low power rectifiers. Packages with glass insulation will be inadequate, as they cannot withstand 400°C temperatures for prolonged periods of time. As a result, the special package shown in Figure 9 was designed. This package utilizes a high percentage alumina ceramic insulator§ and employs a nickel to ceramic brazed joint capable of withstanding temperatures of 600°C.

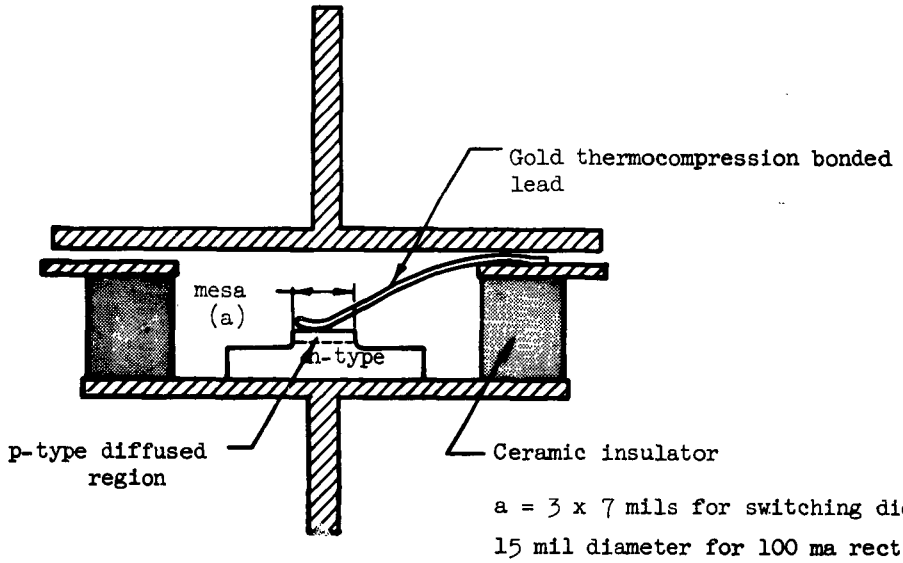


FIGURE 9 SWITCHING DIODE AND RECTIFIER STRUCTURES

2. Device Specifications

The original specifications for a general purpose diode, while realistic at the inception of the contract, were made obsolete due to the rapid strides made in the production of silicon and germanium diodes. As a result of discussions between RCA and the Air Force during the last quarter of 1960, alternative device goals were set:

§Similar to Frenchtown No.4462 ceramic

The objective specifications of the switching diode are:

Peak Inverse Voltage	40 Volts at 300°C Ambient
Leakage Current at 20 volts	100 μa at 300°C Ambient
Average Rectified Current	20 ma at 300°C Ambient
Forward Voltage Drop (20ma)	1 volt at 300°C Ambient
Rectification Efficiency	40% at 100°C Ambient
Recovery Time	1 mμsec at 100°C Ambient
Maximum Operating Temperature	400 °C

The objective specifications of the 100-ma rectifier are:

Peak Inverse Voltage	80 Volts at 300°C Ambient
Leakage Current	500 μa at 300°C Ambient
Forward Voltage Drop (100 ma)	1.3 Volts at 300°C Ambient
Average Rectified Current	100 ma at 300°C Ambient
Maximum Operating Temperature	400°C

The objective specifications of the 250-ma rectifier are:

Peak Inverse Voltage	80 Volts at 300°C Ambient
Leakage Current	500 μa at 300°C Ambient
Forward Voltage Drop (100ma)	1.3 Volts at 300°C Ambient
Average Rectified Current	250 ma at 300°C Ambient
Maximum Operating Temperature	400°C

3. Device Processing

The processing is essentially the same for the switching diode, the 100-ma rectifier and the 250-ma rectifier. Differences between the devices will be pointed out where pertinent.

N-type gallium arsenide crystal with carrier concentrations between 1 and 3 x 10^{16} carriers/cm^3 and room temperature mobility of 4000-5550 cm^2/volt-sec is used for the fabrication of the diodes. It is desirable that the minority carrier mobility at liquid nitrogen temperature be appreciably larger than that at room temperature. Material with carrier concentrations lower than 1 x 10^{16} carriers/cm^3 converted to the non-usable high resistance type gallium arsenide at processing temperatures as low as 650°C.

Suitable crystal is x-ray oriented to the (111) plane and cut into 20 to 25-mil thick slices. After lapping, prediffusion etching and post-diffusion lapping the crystal is 10 to 14-mil thick.

The pre-diffusion etch consists of 20 parts, 5 per cent sodium hydroxide and 4 parts, 30 per cent hydrogen peroxide at the boiling point. The advantage of this etch is that its use will aid in differentiating between the (111) plane and ($\bar{1}\bar{1}\bar{1}$)plane. The (111) plane will appear bright and shiny with coarse pits and the ($\bar{1}\bar{1}\bar{1}$) plane will appear smooth and fine-grained.

Manganese is used as the p-type impurity source, and is evaporated onto the ($\bar{1}\bar{1}\bar{1}$) plane of the wafer. Our experience has indicated that this plane gives higher breakdown voltages than evaporation on the (111) plane. The layer thickness is about 1000Å and is monitored by light transmission on a glass slide.

For diffusion purposes, the manganese coated wafers are placed in quartz tubes together with one piece of oxygen gettering tantalum metal and evacuated to 5×10^{-6} mm Hg in a manifold vacuum system. The quartz tubes are then isolated from the vacuum pump and dry tank nitrogen is backfilled to one-quarter atmosphere. The quartz tubes are sealed off by a stainless steel tipped oxygen-hydrogen torch (to avoid ampoule contamination by copper for the usually employed brass tip).

Diffusions are made between 800 and 1050°C depending on the individual carrier concentration of the wafers. The lower the carrier concentration the lower the temperature (and the longer the time) has to be to avoid thermal conversion to high-resistivity gallium arsenide material. The goal of this diffusion is to obtain a junction depth 3 mils below the crystal surface and a somewhat compensated bulk carrier concentration of about 1×10^{16} carrier/cm^3.

After diffusion the (111) face is lapped off to expose the original n-type material. The reason for this is that surface diffusion and migration will make the (111) face p-type. Tin is then evaporated onto the (111) face and sintered in for 5 minutes in a hydrogen atmosphere at a temperature between 650 and 680°C. The wafers are then electroplated with nickel on both faces.

The mesa area is determined by "smearing" dissolved apiezon wax through a metal alignment mask. For the switching diode the mask has 5 x 9-mil holes. The mesa area is then reduced by undercutting to 3 x 7 mils. For the 100-ma rectifier a 15-mil diameter mesa is used. For the 250-ma rectifier a 30-mil diameter mesa is used.

After mesa formation, the pellets are diced and pre-tested. Good pellets are alloyed into the special package shown in Figure 9. A silver-gold-germanium ternary alloy is used for this purpose. The thermal cycle is the same as for the tin sintering.

Contact between the mesa and the package flange is made by thermocompression bonding a 2-mil wire to each point. The bonding temperature is 480°C. Care must be taken not to punch through the nickel layer which would bring gold in contact with the gallium arsenide. Since the gold-gallium arsenide eutectic point is between 350 and 400°C the gold wire will be dissolved resulting in an open circuit.

An electrolytic clean-up etch in KOH solution is used to improve the reverse characteristic. The voltage is applied in pulses in order to avoid electrolytic polarization. A weak solution of acetic acid is used to neutralize the remaining KOH residue after several water rinsing steps.

Finally the switching diodes are sealed with a 30KVA welder.

4. Device Performance

A group of ten samples each for the three devices described in this section were delivered during the last quarter of 1960. Tables V, VI, and VII give the characteristics of the devices.

As can be seen from Tables V, VI and VII the samples delivered under this contract come very close to the device objectives. It should be noted that the switching diodes are extremely good varactor diodes due to small capacitance and low series resistance.

For high-temperature operation the units have to be derated to avoid thermal run away. In Figure 10 derating curves for different dynamic forward resistance are given; forty-five per cent of the 250-ma rectifier have a dynamic resistance equal or less than one ohm, another 40 per cent fall between 1 and 2 ohms. Units with less than 1 ohm can be operated to full 250-ma rectified current at 300°C while units with 2 ohms can be operated only to 150-ma rectified current.

The average thermal resistance for the 250-ma rectifiers is 85°C/watt, and can be reduced by heat sinking the case.

The reverse current is a function of temperature. In Figure 11 reverse current for the 250-ma rectifiers is plotted in a statistical manner. It can be seen that at higher temperatures the reverse current distribution is more normal (thermal current across the gap) while at lower temperatures spurious leakage is more predominant.

Figure 12 is a replot of the mean 250-ma unit. The reverse current as a function of temperature has an activation energy of $E = 1.05$ ev. This is quite similar to silicon devices which show activation energy levels of 0.7 ev. instead of the band gap of 1.1 ev.

TABLE V

CHARACTERISTICS OF GALLIUM ARSENIDE SWITCHING DIODES

Diode No.	Forward Voltage Drop, volts I_f = 20ma, T = 25°C	Peak Inverse Voltage, volt I_r = 10μa T = 25°C	Reverse Current, μa, V_r=.6V_B I_f = 20ma T = 300°C	Dynamic Resistance, ohms T = 25°C	Rectification Efficiency, % 100mc = 2vrms T = 25°C	Recovery Time, n sec. I_f=10ma, V_r= 6 volts, T = 25°C	Capacitance, pf. V_r = 1.5 volts
1	1.2	61	45	15.0	46.1	1.4	4.80
4	1.3	62	20	11.5	44.7	1.3	7.40
7	1.3	46	24	12.5	48.6	1.0	4.60
15	1.4	76	25	17.5	45.1	1.4	2.36
17	1.3	52	17	13.8	46.8	1.3	2.46
18	1.8	71	200	22.5	45.4	1.2	1.81
20	1.4	59	50	23.7	43.3	1.3	1.90
21	1.3	60	30	16.2	41.8	1.7	2.21
22	1.3	54	22	11.2	44.4	1.6	3.20

TABLE VI

CHARACTERISTICS OF 100-ma GALLIUM ARSENIDE RECTIFIERS

Diode No.	Breakdown Voltage, volts I_r=10μa, T=25°C	Forward Voltage, volts I_f=100ma, T=25°C	Reverse Current, ma I_f=100ma, V_r=.6V_B T = 300°C	Forward Resistance, ohms
1	100	2.15	130	8.4
3	108	2.50	150	10.0
4	98	2.10	190	8.0
5	172	2.90	110	10.0
6	148	3.30	52	14.0
7	97	2.10	75	9.0
8	96	2.30	50	9.0
9	100	2.35	42	11.0
12	84	2.50	180	12.0
13	82	3.80	77	10.0

TABLE VII

CHARACTERISTICS OF 250-ma GALLIUM ARSENIDE RECTIFIERS

Diode No.	Breakdown Voltage, volts I_r=10 μa T=25°C	Forward Voltage, volts I_f=10ma, T=25°C	Forward Voltage, volts I_f=250ma, T=25°C	Reverse Current, μa T=300°C, V_r = .6V_B
1	72	1.1	1.6	360
2	72	1.2	1.7	140
3	70	1.1	1.5	220
4	70	1.1	1.5	300
5	80	1.1	1.9	320
6	76	1.2	3.0	320
7	64	1.0	1.5	180
8	68	1.1	1.7	250
9	64	1.1	1.8	200
10	68	1.1	1.9	165

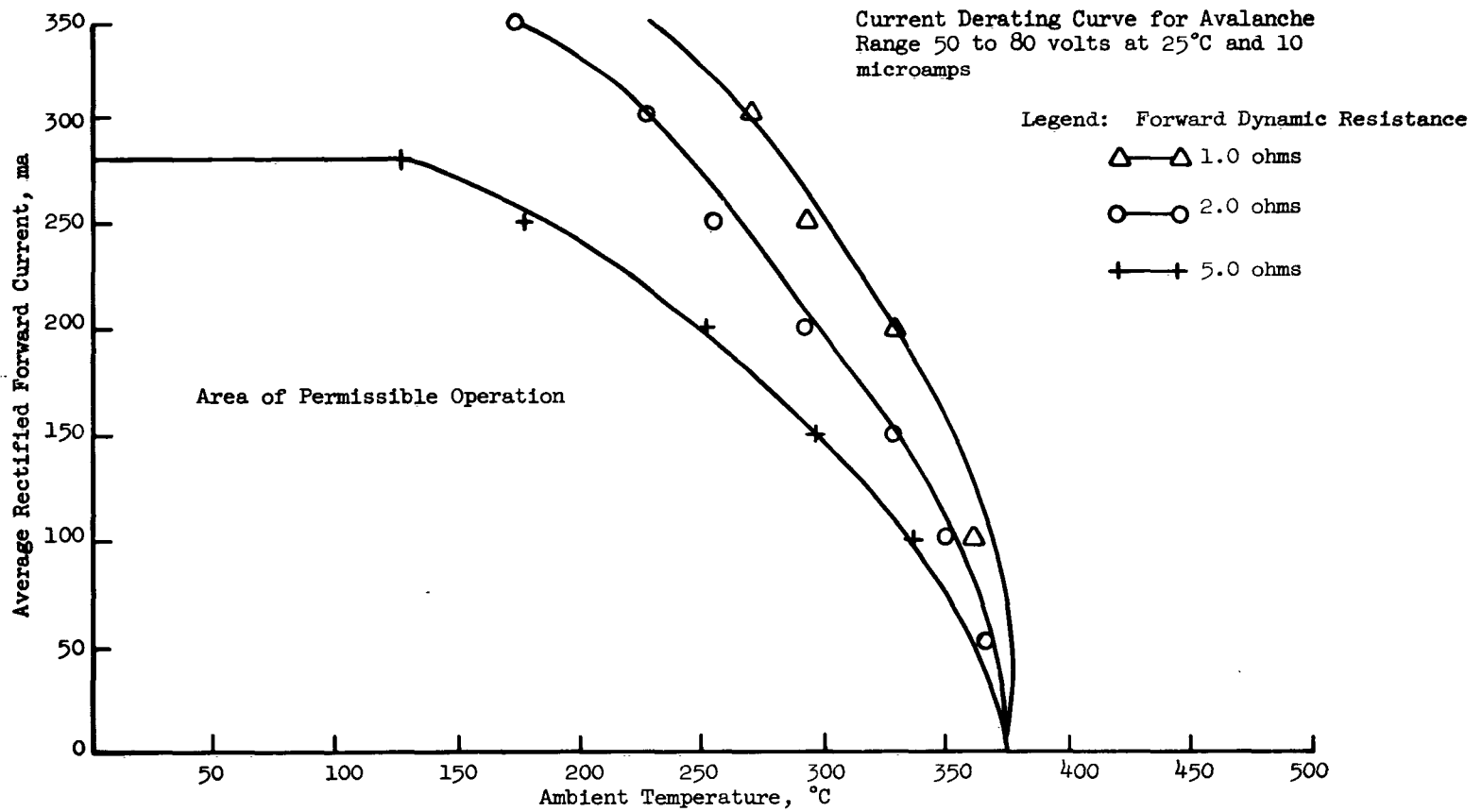


FIGURE 10 CURRENT DERATING CURVE FOR 250-ma GALLIUM ARSENIDE RECTIFIERS

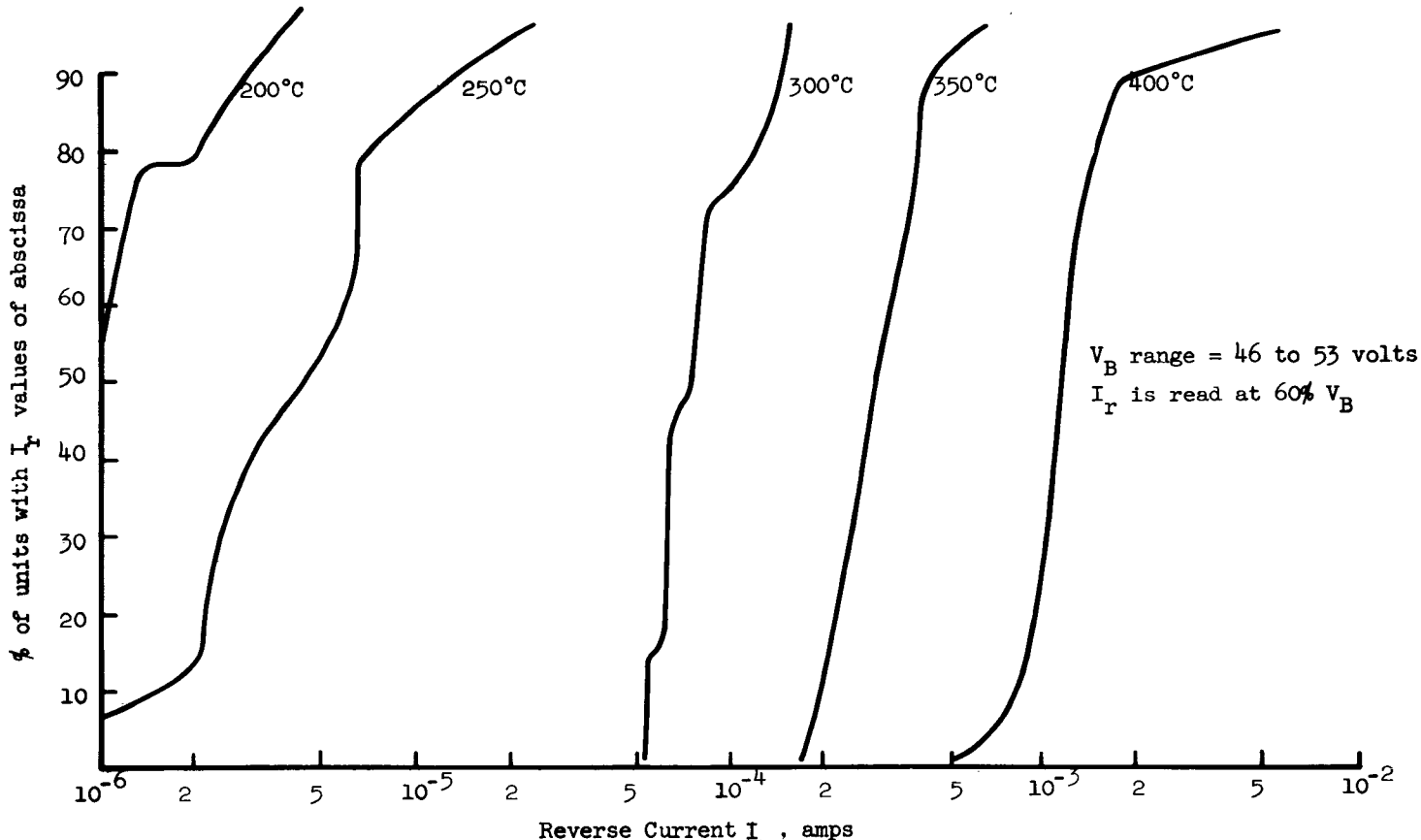


FIGURE 11 REVERSE CURRENT IN 250-ma GALLIUM ARSENIDE RECTIFIERS

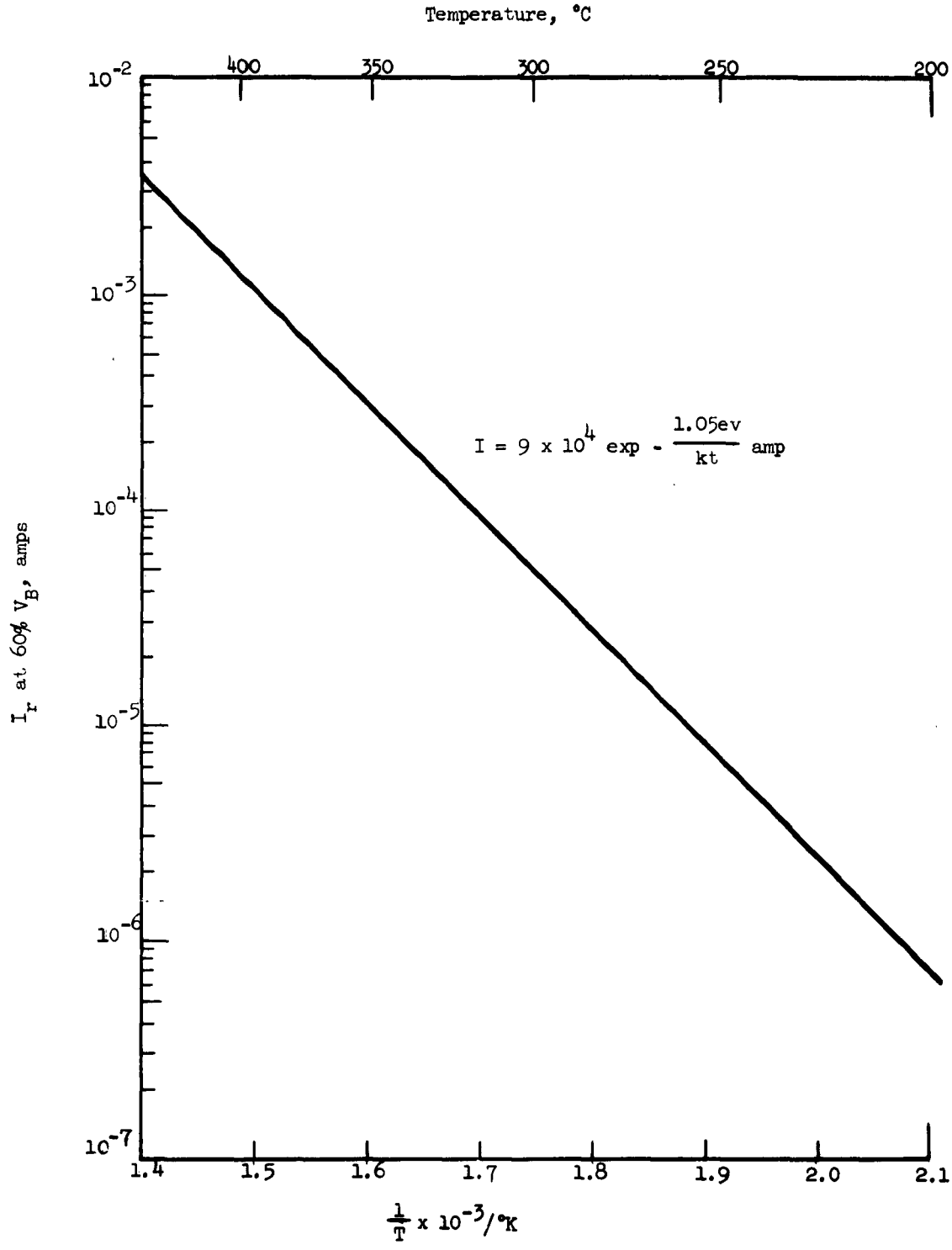


FIGURE 12 REPLOT OF THE MEAN UNIT OF FIGURE 11

D. Varactor Diode

No work has been done on varactor diodes under this contract, but RCA has carried out extensive investigations in this area under other government and RCA sponsored programs. One such program was successfully completed when RCA developed a varactor diode with higher cutoff frequencies then obtainable with silicon varactors under Contract No. AF33(616)-6724 for Wright-Patterson Air Force Base. The diode was used in a subsequent phase of the same contract to produce a very low noise X-band parametric amplifier. In addition to this contract, RCA has recently received a contract from the Department of the Navy, Bureau of Ships to produce varactor diodes with power outputs of up to 100mw at 25°C. RCA also has commercial programs for the development of higher-cutoff-frequency, higher power varactors.

E. Zener Diodes

1. Design Considerations

In the last few years "Zener" or voltage regulator diodes have obtained a position of prominance in electronic circuits. Even though Zener, in 1934, predicted the field emission breakdown effect (tunneling of electrons from the valence band under a strong electric field to the conduction band), most of the present commercial "Zener" diodes are based on the avalanche breakdown phenomenon. The difference in the two phenomena is that the field emission effect causes a negative temperature coefficient due to the decreasing band gap with increasing temperature and the avalanche mechanism shows a positive temperature coefficient due to the increased number of collisions with the crystal lattice with increasing temperatures.

The most desirable Zener diodes have a zero temperature coefficient in the transition region between field emission and avalanche. In this region, positive and negative temperature coefficients just compensate each other. While it is theoretically possible to change concentration gradients, in practice the voltage range for zero temperature coefficients is rather narrow. For silicon Zener diodes the range is around 4 volts and for gallium arsenide zener diodes it is about 6 volts. This higher value obtainable with gallium arsenide is more desirable for most applications of zener diodes.

Unfortunately, not enough data on the temperature coefficients of the two effects are available to theoretically design Zener diodes. However, Figure 13 gives a good indication of what can be expected. In Figure 13 silicon and gallium arsenide Zener diodes have been compared. It will be noted that at higher voltages there is little difference between the two. At lower voltages gallium arsenide diodes have the same temperature

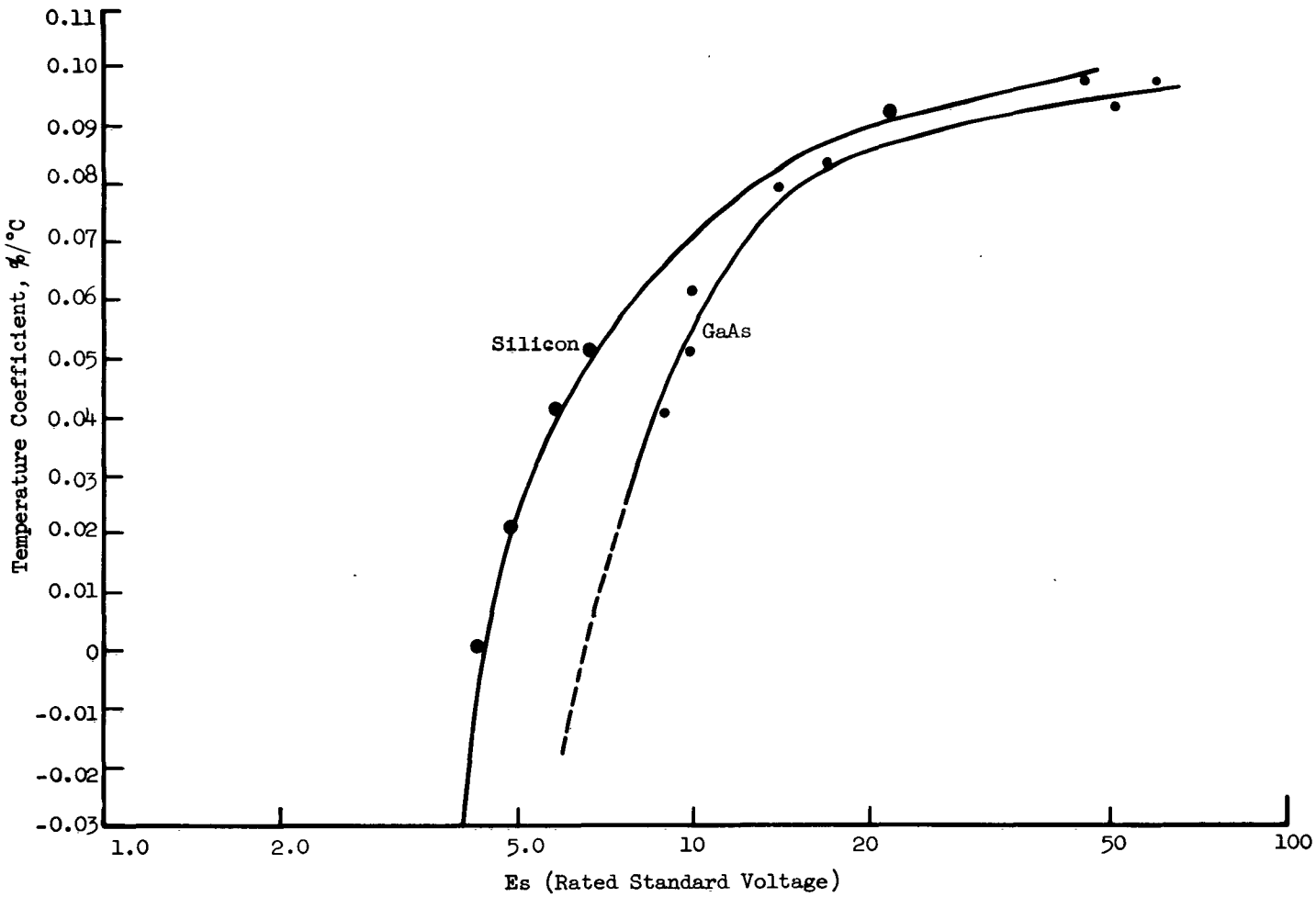


FIGURE 13 REVERSE BIASED AVALANCHE VOLTAGE TEMPERATURE COEFFICIENT VERSUS 25°C AVALANCHE VOLTAGE

coefficients as silicon diodes for somewhat higher values. In the last quarter of 1960 highly doped diodes with low breakdown voltages were made. Unfortunately, the diodes produced did not yield voltage breakdown below 9 volts, therefore, the gallium arsenide curve can only be dashed below this value.

Dynamic impedance tests at 1000 cycles in the breakdown region were made on those diodes using a Boonton Radio 275A transistor test set. The gallium arsenide devices were fully comparable in performance with silicon units.

A typical plot for a gallium arsenide diode is shown in Figure 14. The straight lines show the breakdown slope versus temperatures and reverse current. The curve is a plot of dynamic impedance at 1000 cycles and 25°C and is comparable to a silicon diode in the same voltage category.

2. Device Processing

Except for more highly doped gallium arsenide material the processing is identical to that of the gallium arsenide switching diodes.

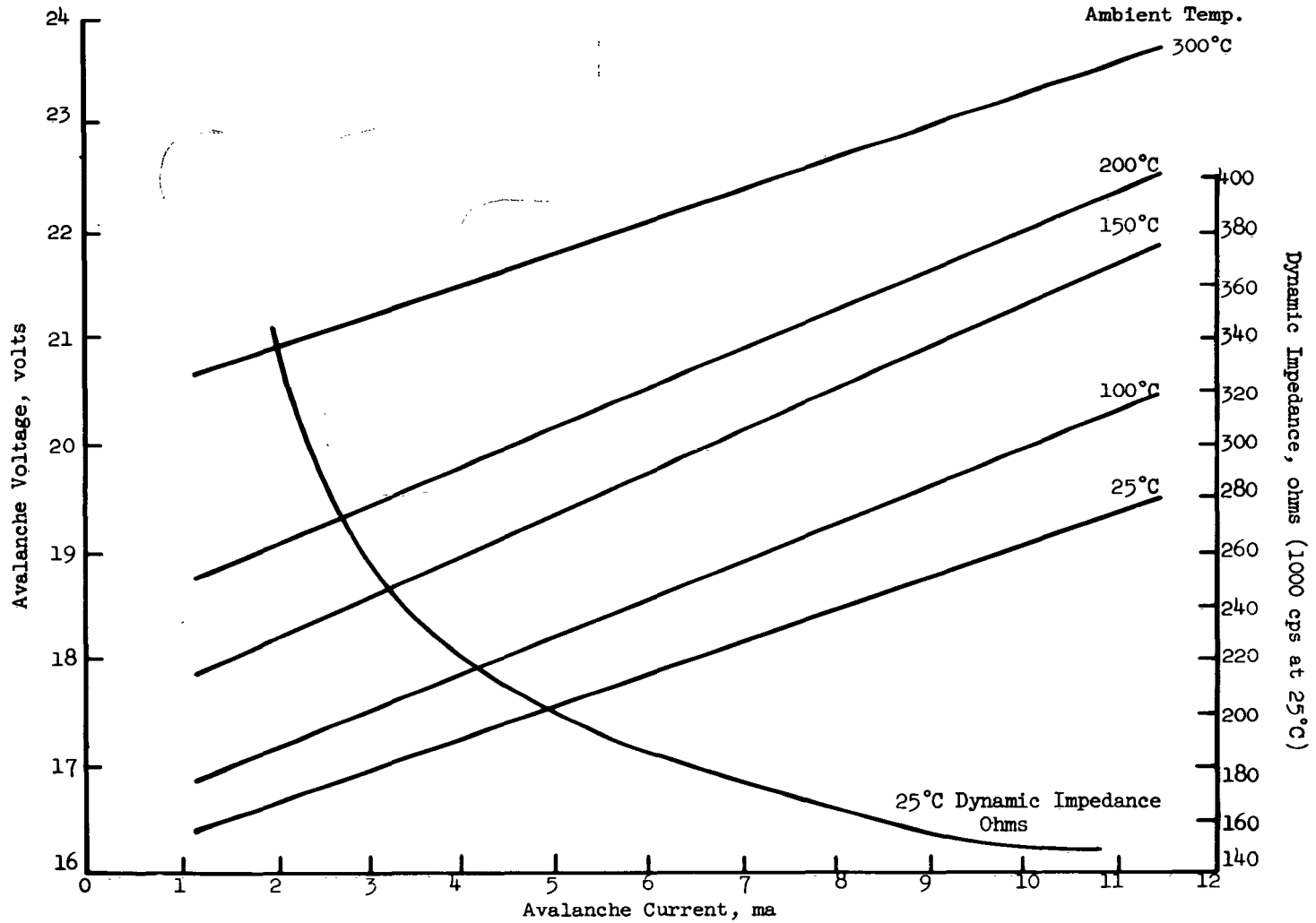


FIGURE 14 AVALANCHE CURRENT VERSUS AVALANCHE VOLTAGE AND DYNAMIC IMPEDANCE

F. Tunnel Diodes

From July to November, 1959 effort was directed toward fabricating the first existing gallium arsenide tunnel diodes. The basic technology used for these samples is still being used today.

Highly doped p-type gallium arsenide is either grown from the melt or made by zinc diffusion and the tunnel junction formed by alloying tin dots to the wafer. After alloying the small gallium arsenide pellets are mounted in pill box packages.

The electrical characteristics of the gallium arsenide tunnel diodes produced were very encouraging. Peak to valley current ratios of up to 45 had been observed with lead inductances as low as 400ph. Basic circuit concepts have also been conceived for application of the tunnel diodes as switches, oscillator amplifiers and down converters.

Since this area looked so promising, follow-up contracts were awarded to RCA. Under Contract NObsr-77523 high-speed switching applications were investigated for a 1-kmc computer. Under contract AF33(616)-7314 tuneable oscillators operating near 1000 and 10,000 mc were investigated as were voltage tuneable oscillators. In addition, low noise microwave amplifiers, power microwave amplifiers, and frequency converters were studied. More detailed information has been given in ASD Technical Report 61-496, pt. I.

G. Power Rectifier

1. Design Considerations

The power rectifier was designed primarily for use at an ambient temperature of 250°C with a maximum storage or junction temperature of 400°C. Selection of the 250°C temperature was arbitrary but the 400°C maximum storage or junction temperature reflects the limit imposed on the gallium arsenide by the band gap energy. As a result, all materials and processes used in making the power rectifier must be consistent with these temperature limitations. The electrical characteristics desired for the power rectifier are:

Peak Inverse Voltage	80 Volts at 250°C Case Temperature
Forward Voltage Drop at 2A	2.5 Volts at 250°C Case Temperature
Leakage Current	500μa at 250°C Case Temperature
Avg. Rectified Current	8 Amperes at 250°C Case Temperature
Max. Operating Temperature	400°C

The power rectifier is a large area pn junction mounted in a suitable package. The basic problems involved in making the rectifier can be broken down into a number of more or less independent general areas: 1.) design and construction of a suitable package, 2.) formation of the pn junction, 3.) mounting and contacting and 4.) stabilization and sealing. Although there is some overlap between the various areas, they can generally be considered as separate problems.

In designing a package for the power rectifier, the major considerations are 1.) operating temperature limitations, 2.) power dissipation requirements and 3.) electrical resistance of the contacts and leads. Because of the 400°C temperature requirement, a metal-ceramic structure was selected. Figure 15 shows the construction of the package and also how the rectifier is mounted into this package. The large copper base is 0.625 inch in diameter and 0.187-inch high and serves as the primary heat conduction path.

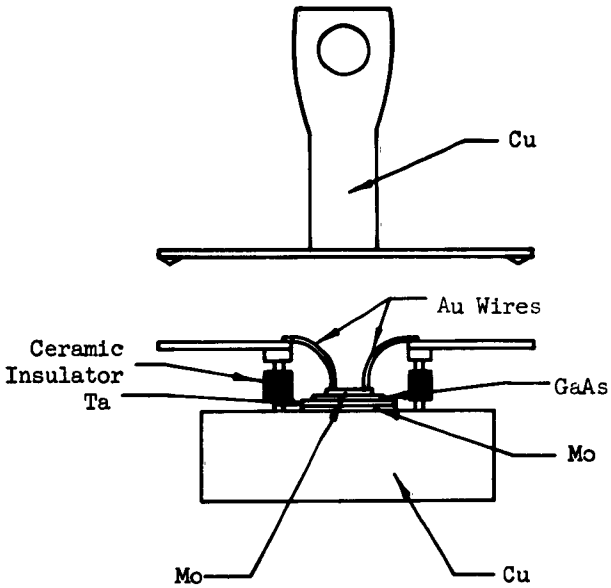


FIGURE 15 ASSEMBLED POWER RECTIFIER

The temperature rise for the package with a rectifier mounted as shown is about 2.5°C per watt of dissipation. The package is hermetically sealed by ring welding. To prevent oxidation at high operating temperatures, all metal parts of the package are nickel plated.

The rectifier pellet is a diffused junction structure 180 mils in diameter with a 135-mil diameter mesa. To arrive at a design adequate to meet the proposed requirements, a number of compromises were necessary in designing the junction. The maximum breakdown voltage that can be obtained in a design of this type is about 80-90 volts. This is due to the thermal conversion problems encountered during diffusion of high resistivity gallium arsenide. At the same time, conductivity modulation does not occur to any significant extent to reduce the voltage drop in the forward direction. It would be possible to reduce the forward voltage

drop by making the junction area very large, but this would also make the reverse leakage too high. Therefore, the 135-mil junction on a 180-mil pellet is a compromise designed to give acceptable values for all parameter.

Primary design considerations in assembly were to maintain a 400°C temperature capability and to provide low thermal and electrical resistance. The 400°C temperature limit requires the use of high-temperature solders which are characteristically quite hard. It was, therefore, necessary to develop a mechanical system that would not crack the gallium arsenide pellet during soldering.

2. Device Processing

The starting material for the power rectifier is n-type gallium arsenide with a carrier concentration of from 1 to 4 x 10^{16} carriers/cm^3. The pn junction is formed by a deep manganese diffusion which results in a shallow gradient that yields a high breakdown voltage. The diffusion is done at 1050°C for 40 minutes in a sealed quartz ampoule. The manganese source is evaporated onto the back of the wafer prior to diffusion. A major problem at this step in the process was the partial or complete conversion from n-type to p-type during the high-temperature diffusion. The tendency to convert was minimized by using the evaporated manganese source, by very careful cleaning of the ampoule and wafers prior to ampoule sealing, and by back-filling ampoules to about one-half atmosphere of pressure with either nitrogen or forming gas. A number of other experiments such as radiant furnace diffusion, zinc diffusion and various gettering processes were tried but none of these techniques were completely successful. The thermal conversion of all high-resistivity gallium arsenide represents a fundamental limit on the breakdown voltage that can be achieved with rectifiers.

After diffusion, one side of the gallium arsenide wafer is lapped to remove the p-type layer and both sides of the wafer are nickel plated. Circular pellets are then formed by masking the wafer. The pellets are separated by sandblasting. The p-type side of the pellet is then masked with a 125-mil disc and the nickel removed. The pellets are etched electrolytically in potassium hydroxide to remove the gallium arsenide damaged by the sandblasting.

Mounting is also a fairly complex procedure because of the 400°C temperature requirement. The gallium arsenide pellet cannot be soldered directly to the copper package because of thermal mismatch between gallium arsenide and copper. The problem of pellet cracking due to the thermal mismatch was minimized by the use of the construction shown in Figure 15. In this construction, a molybdenum disc is brazed into the package and a tantalum disc is brazed on top of the molybdenum; both operations use RTSN brazing alloy. Next the gallium arsenide is soldered to the tantalum using an alloy of gold, silver and germanium doped with palladium. At the same time a molybdenum disc is soldered to the p-type side of the gallium arsenide pellet using an alloy of gold, silver, and germanium doped with zinc. This sandwich construction using molybdenum and tantalum discs with the high-temperature solders was developed to prevent cracking of the gallium arsenide pellets during mounting and any subsequent heat treatments. Contact to the top molybdenum disc is made by nailhead bonding five or six five-mil gold wires to the molybdenum disc and the top flange of the package.

After contacting, the units are etched in potassium hydroxide solution to improve the electrical characteristic of the junction. It was found that the etch pattern could be directed to the pn-junction using a high current pulse etch and light illumination.

The rectifiers are next baked out at 250°C in a forming gas atmosphere and then sealed in dry nitrogen. The forming gas atmosphere is used to prevent degradient of the rectifier when they are baked at temperatures greater than 200°C.

3. Device Evaluation

A variety of measurements have been made on the power rectifiers. Figure 16 shows the results of measurements of reverse leakage current as a function of temperature. Figure 17 shows reverse leakage currents as a function of voltage. Measurements of forward resistance on several rectifiers have indicated a range of values from 0.04 ohms to 0.2 ohms with an average of about 0.08 ohms. Thermal resistance measurements show most units to be in the range between 2.3 and 3.0°C per watt dissipation. Some additional data on the 10 state-of-the-art samples submitted to the Air Force is given in Table VIII.

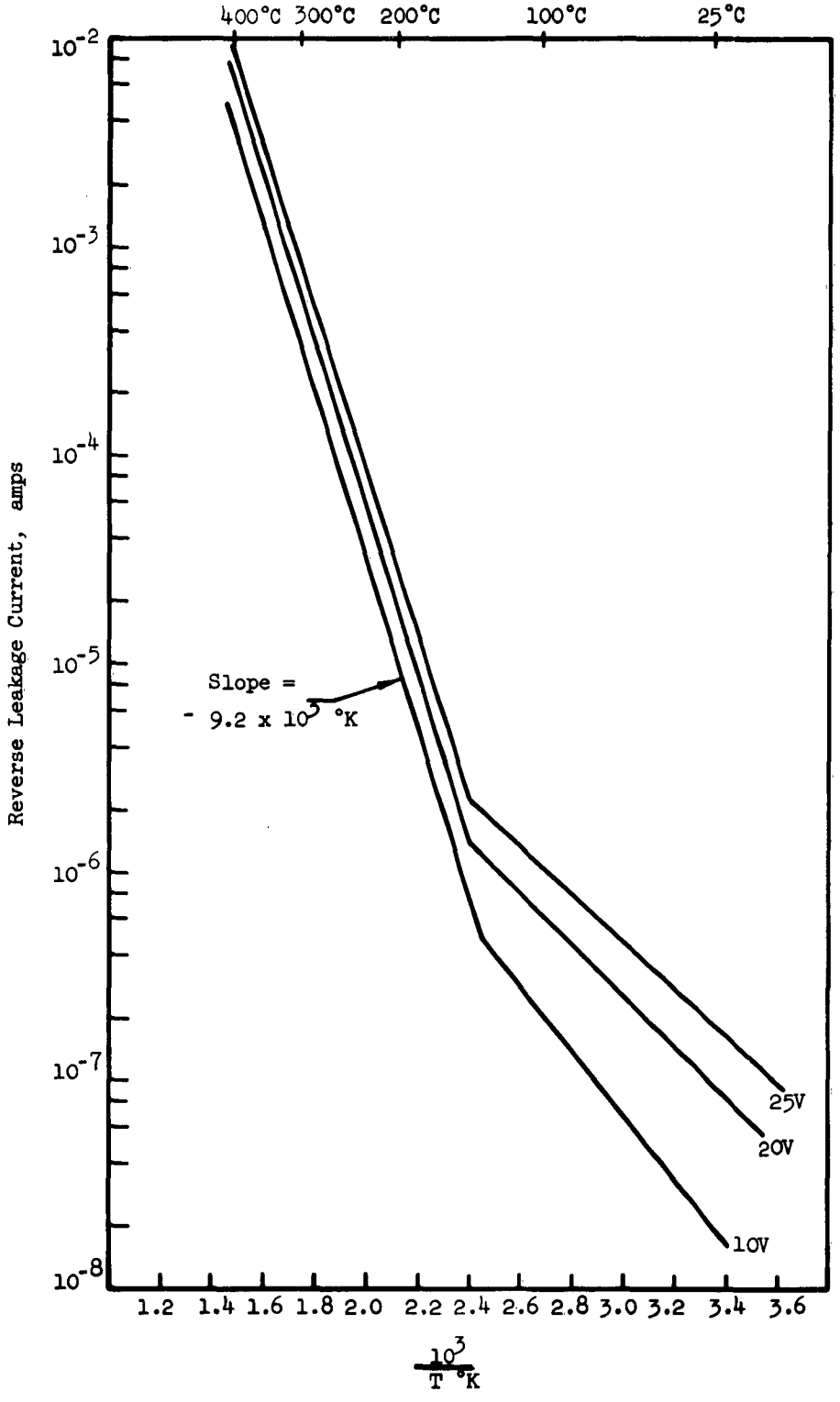


FIGURE 16 REVERSE LEAKAGE CURRENT VERSUS $\frac{1}{T\ °K}$ FOR SEVERAL REVERSE VOLTAGES

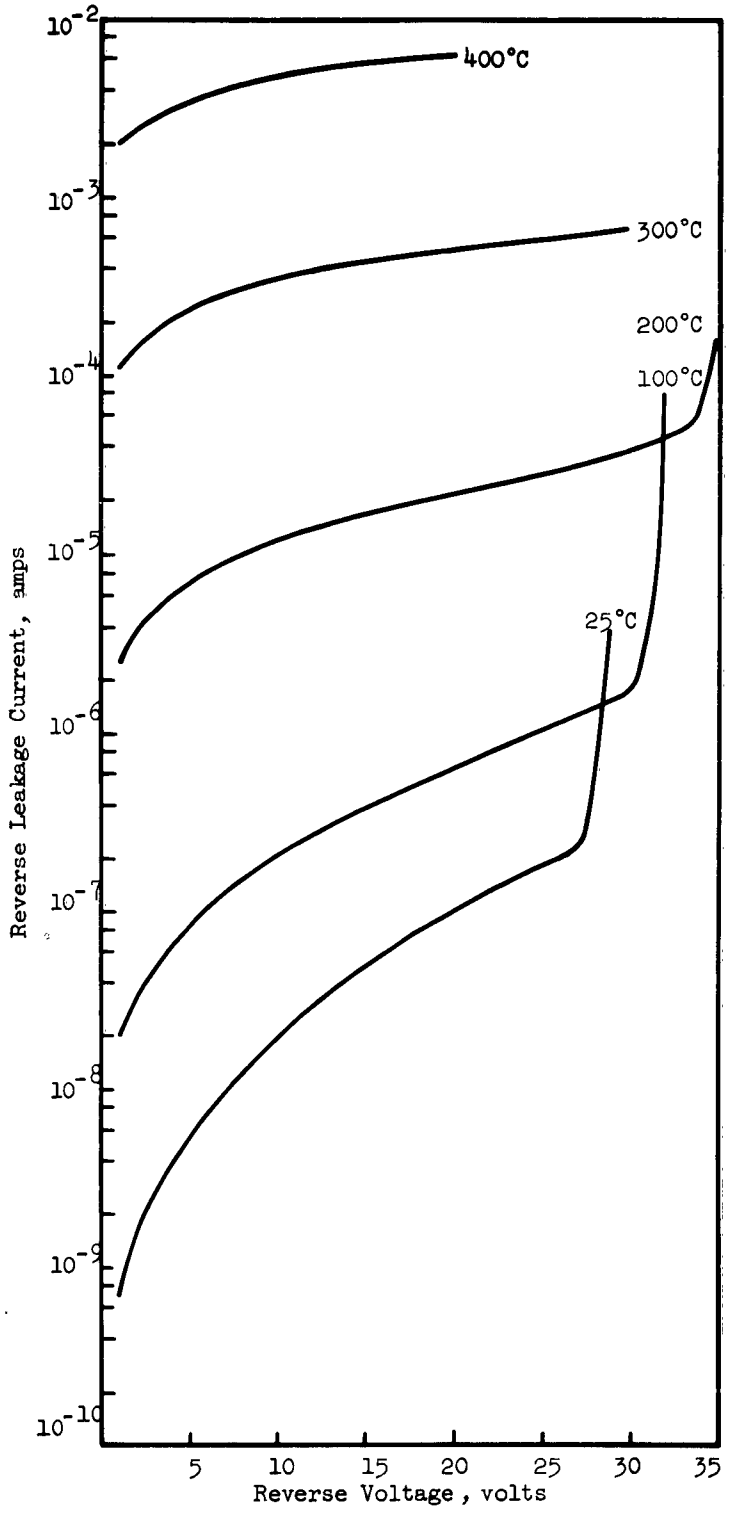


FIGURE 17 REVERSE LEAKAGE CURRENT VERSUS REVERSE VOLTAGE AT VARIOUS TEMPERATURES

TABLE VIII

DATA ON 10 STATE-OF-THE-ART POWER RECTIFIERS

Unit No.	PRV, volts	V_f, volts	I_r, μa T = 250°C
1	64	1.7	580
2	73	1.8	230
3	77	1.4	820
4	67	1.4	1600
5	73	1.5	1320
6	47	1.9	440
7	54	1.7	800
9	68	1.35	330
13	53	1.7	820
15	69	2.7	1050

PRV is the reverse voltage at which the reverse current reached 50μa at room temperature.

V_f is the forward voltage for 2 amperes of forward current flow at room temperature.

I_r is the reverse current at 250°C and 60% of PRV. There was no forward current during this particular measurement.

H. Solar Cells

Although no work on gallium arsenide solar cells was carried out under this contract, in 1959, the Semiconductor and Materials Division of RCA undertook a Research and Development program on gallium arsenide solar cells under Air Force Contract Number AF33(616)-6615. Using the base established at the RCA Laboratories under Signal Corps Contract DA-36-039-sc-78184, cells with efficiencies of eight to eleven per cent were obtained routinely with higher efficiencies obtained on a number of occasions. In addition, gallium arsenide solar cells were found to be superior to other cells in resistance to radiation of the type found in the Van Allen Belt. Gallium arsenide cells are also superior in the reduction of efficiency suffered with increased temperature operation.

In the summer of 1961, NASA became interested in gallium arsenide solar cells because of their high resistance to radiation. RCA, Somerville was given a contract, NASA contract Number NAS5-457 to produce 200 gallium arsenide solar cells for evaluation. Some of these cells were mounted on the first "RELAY" satallite, built by the ASTRO-Electronics Division of RCA.

Presently gallium arsenide solar cells are being produced in a pilot line operation and are being evaluated by a number of organizations throughout the country.

I. Unipolar Transistors

1. General

Although the principles of field effect (unipolar) action were known (ref.6) long before those of the bipolar transistor action, very little effort has been devoted to the field effect transistor.

Certainly one reason for lack of effort on the field effect transistor was, at least originally, the relatively high amplifier noise figure (ref.7). Another reason was the almost hundredfold lower transconductance compared to the bipolar transistor which, in semiconductor device history, (ref.8,9) was observed very early. Recent analyses (ref.10,11) on quite general grounds have shown that the field effect transistor is inherently inferior with respect to transconductance, and, thus, probably inferior with respect to the all-important gain-bandwidth product.

Lately (1961), however, the interest in germanium and silicon unipolar transistors has increased. This is due to the reduced noise figure which was obtained with the planar device technology. As a matter of fact the noise figures with this type of device are lower than either bipolar transistors or vacuum tubes. In addition, photolithographic processing allows

---

ref. 6. O. Heil, British Patent 439, 457; 1935

ref. 7. Discussions with the Bell Laboratories' staff.

ref. 8. W. Shockley, "A Unipolar 'Field Effect' Transistor", Proc. I.R.E., 40, pp 1365-76; November 1952

ref. 9. G.C. Dacey and I.M. Ross, "The Field Effect Transistor", Bell Sys. Tech. Jour., 34,pp. 1149-1189; November 1955.

ref.10. J.M. Early, "Structure-Determined Gain Band Product of Junction Triode Transistors" Proc. I.R.E., 46, pp.1924-27, December 1958

ref.11. E.O. Johnson and A. Rose, "Simple General Analysis of Amplifier Devices with Emitter, Control, and Collector Functions", Proc. I.R.E., pp.407-418, March 1959.

extremely small dimensions and, therefore, respectable frequency performance.

Gallium arsenide is a better unipolar transistor material than either germanium or silicon. (See Section A). Furthermore, since minority carrier lifetime is of no importance to unipolar action there is fundamentally a better chance of obtaining power gain with gallium arsenide unipolar transistors than with gallium arsenide bipolar transistors on low lifetime material.

From July to November 1959 an active program on gallium arsenide unipolar transistors was undertaken. Since the basic technology was not sufficiently developed relatively crude gallium arsenide unipolar transistors were fabricated.

2. Design Consideration

The unipolar transistor is basically a modulated resistor. In its simplest form, it consists of an extrinsic semiconductor bar with ohmic contacts affixed to both ends. One contact is termed the source where majority carriers enter the device, and the other is called the drain where the carriers leave the device. Interposed between the source and drain contact areas is a region of the opposite conductivity type called the gate, whose function is to vary the resistance from source to drain. The region beneath this gate is the active portion of the transistor and is called the channel. If a voltage is applied between the channel and the gate, in a direction such as to reverse bias this junction, the depletion region is caused to expand in the channel, thus increasing the resistance of the channel.

To illustrate the operation, a simplified unipolar transistor is sketched in Figure 18. The drain current, I_D, and source-to-drain voltage, V_{SD}, are related through the conductance G[§], of the device.

§To simplify matters, the conductance, G, is assumed to lie totally within the active region of the device (i.e., in the channel).

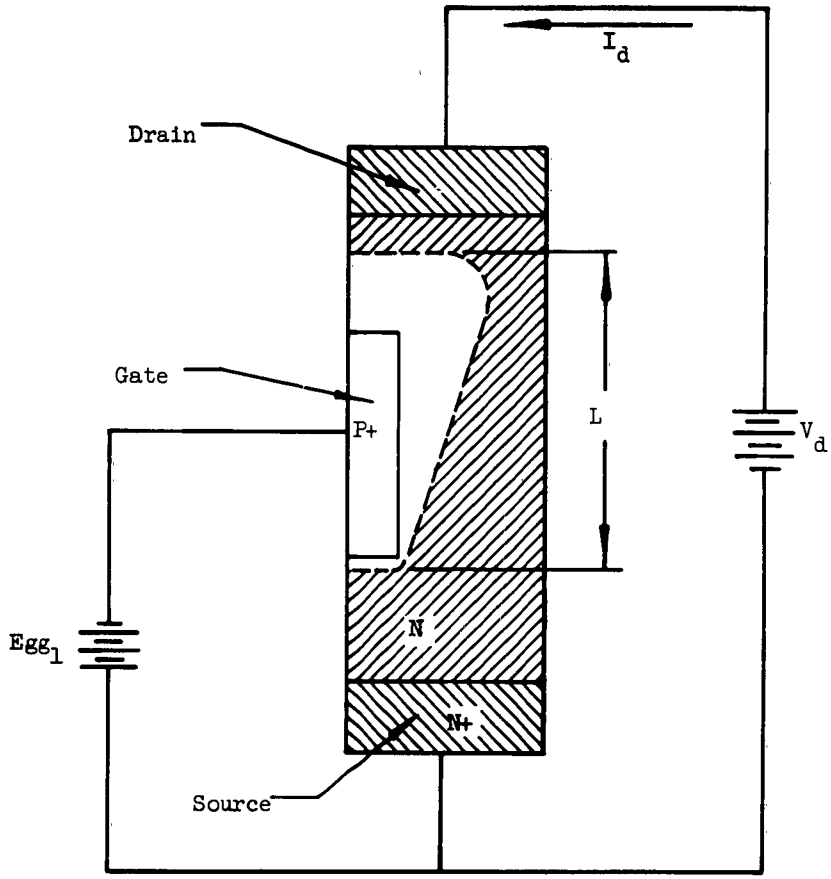


FIGURE 18 SIMPLIFIED UNIPOLAR TRANSISTOR

$$I_D = GV_D \tag{5}$$

The conductance of the channel may be expressed in more fundamental terms: conductivity, σ, cross section area, A, and channel length, L.

$$G = \frac{\sigma A}{L} \tag{6}$$

Changing any of these parameters changes G, and thus disturbs I_D, V_{SD}, or, more likely, both. The unipolar transistor modulates G by

variation in cross sectional area, A, through the use of the gate depletion layer. A further extension of the expression for G into still more fundamental concepts leads to the charge control concept. If the conductivity σ is re-expressed as the product of mobility, μ, mobile carrier density, N, and electronic charge, g, the following equation results:

$$G = \mu n \left(q \frac{A}{L} \right) \tag{7}$$

If this equation is re-arranged, the conductance G may be expressed as:

$$G = \mu \left(q \frac{ALN}{L^2} \right) \tag{8}$$

with the product (qALN) appearing in the numerator. This product is simply the total mobile charge, Q.

$$G = \mu \left(\frac{Q}{L^2} \right) \tag{9}$$

Examination of Equation (9) indicates that modulation of Q results in modulation of G. From this view, the control electrode (the gate) is simply one plate of a condenser to which mobile charge is transferred and stored, thus being made inactive. These two modulation concepts are essentially the same; however, each has certain merits in explaining facets of unipolar operation.

3. Device Processing

A 20-mil thick wafer of n-type gallium arsenide, as high resistivity as possible ($\sim$0.01Ω-cm) is placed in an evacuated diffusion ampoule with a zinc pellet and small tantalum squares. The zinc is then diffused with gallium arsenide at 1100°C for one-half hour. This operation results in a diffusion depth of 4 to 5 mils.

Next, the p-type skin on one side of the wafer is lapped off to expose the original n-type material. This side of the wafer is lead coated from an evaporated source and then plated with nickel. These materials are sintered in at 600°C.

The n-type side is then masked with wax (see Figure 19) and the gate junction area is defined by etching. Next, the channel location on the n-type side is defined by small pieces of tape. A slow etch (1 mil per minute) is then applied to establish the channel thickness. Etching and measurements of field effect action are carried out alternately. This process is stopped when a pronounced field effect action is observed.

The finished pellet is soldered to a header with lead.

The most critical step in the process is the establishment of the narrow channel. Uneven diffusion of the junction is the main difficulty. If the impurity density in the gallium arsenide were 10^{15} carriers/cm^3 instead of 7 to 8x10^{16} carriers/cm^3, the channel thickness could be very much greater and careful dimensional control would not be necessary. As it is now, a slight miscalculation in the final etching step leads to punch-through and very poor gate junctions and channels.

4. Device Performance

Measurements were taken on 11 unipolar devices. The results of these measurements are presented in Table IX. Next, the devices were measured in the circuit shown in Figure 20. The results of these measurements are shown in Table X. Of first order interest are the transconductances calculated from the observed modulation of the channel current. These have values in the 50 to 300 μmhos range. These values could be increased by increasing the channel bias voltage, V_{Do}, or by decreasing the channel length and/or increasing its lateral dimension. For Unit No. 1 the voltage, V_{Do}, can be increased roughly four-fold if we assume that the unit can tolerate a dissipation of 100mw in the channel. This increase in V_{Do}, bought at the expense of high internal power dissipation, would cause a proportional increase in I_{Do}, ΔI_D, and the transconductance. A better way to improve the performance would be to change

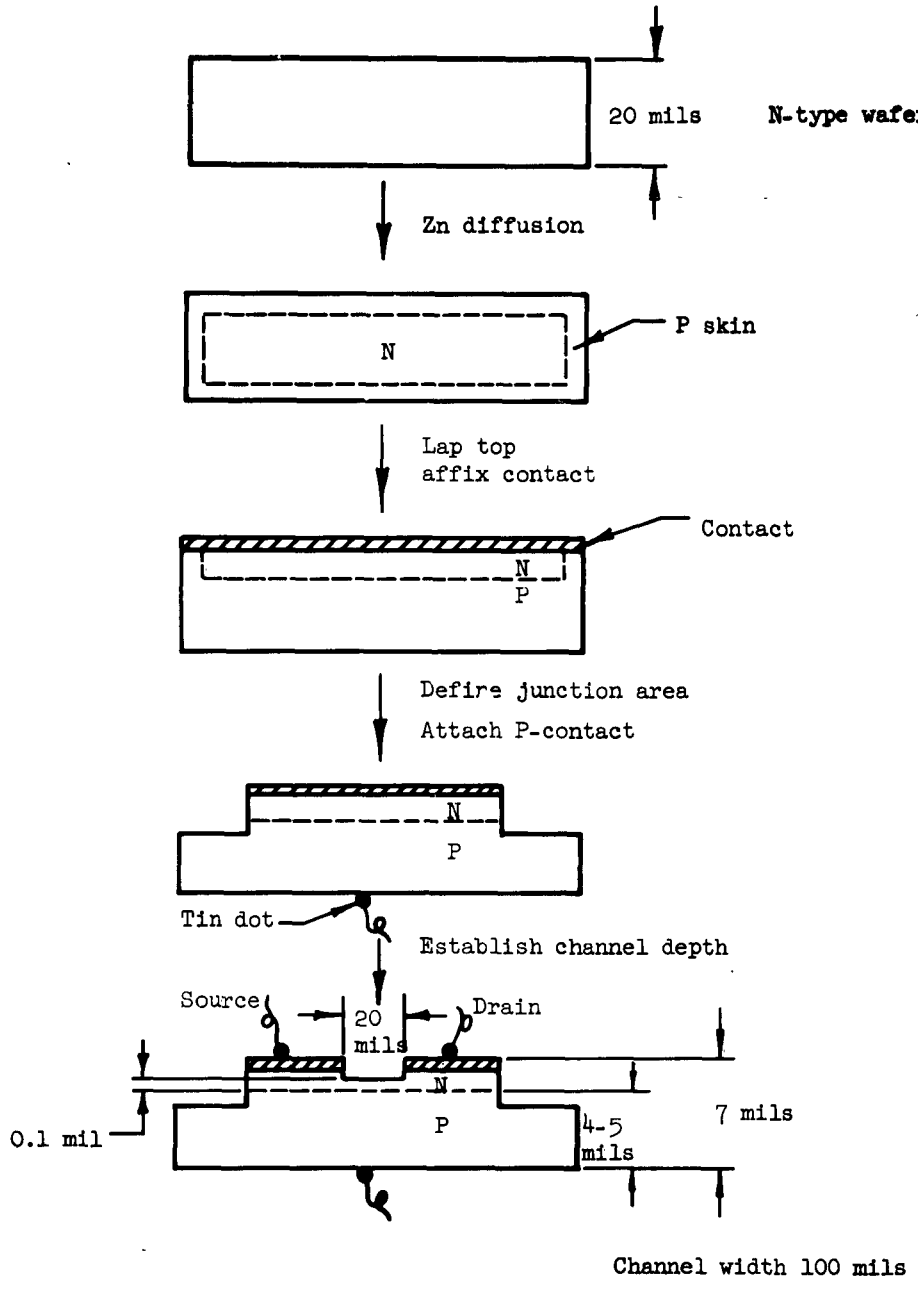


FIGURE 19 FABRICATION PROCESS FOR UNIPOLAR TRANSISTOR

TABLE IX

UNIPOLAR TRANSISTOR DATA

Unit No.	Original Crystal		Gate		Channel	
	n, cm^{-3}	μ, $cm^2\text{-}sec^{-1}\text{-}volt^{-1}$	Max. Voltage, volts	Reverse Current, μa	R_c, Ω	Length(L), cm
1	7×10^{16}	6000	30	~1	100	.05 nominal
2	8×10^{16}	5090	20	<1	96	.05 nominal
3	7×10^{16}	5000	10	<3	165	.05 nominal
4	7×10^{16}	5000	15	<1	124	.05 nominal
5	7×10^{16}	5000	15	<1	318	.05 nominal
6	8×10^{16}	5090	10	<1	510	.05 nominal
7	7×10^{16}	5000	15	<1	80	.05 nominal
8	8×10^{16}	5090	7	<1	580	.05 nominal
9	8×10^{16}	5090	15	<1	112	.05 nominal
10	8×10^{16}	5090	15	<1	220	.05 nominal
11	8×10^{16}	5090	8	<1	124	.05 nominal

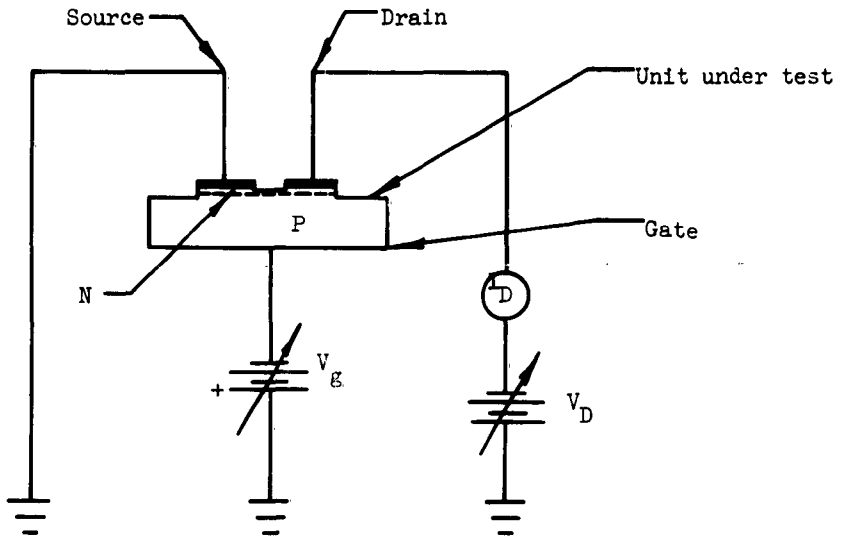


FIGURE 20 UNIPOLAR TRANSISTOR MEASURING CIRCUIT

TABLE X

OBSERVED PERFORMANCE OF UNIPOLAR DEVICES MEASURED IN FIGURE 20

	Bias State				Modulation						
Unit No.	I_{Do}, ma	V_{go}, volts	I_{go}, μa	V_{Do}, volts	ΔV_g, volts	ΔI_g, μa	ΔI_D, ma	$\frac{\Delta I_D}{\Delta V_g}$, μmhos	ΔP_i, watts	ΔP_o*, watts	D.C. Power Gain(G_p)**
1	5	20	1	0.85	20	~1	1.1	50	2.5×10^{-6}	12.5×10^{-6}	5
2	5	10	<1	0.47	10	<1	1.33	1.33	1.25	21.2	17
3	5	10	<3	0.825	10	<3	1.07	107	3.75	23.6	6.3
4	5	10	<1	0.62	10	<1	1.12	112	1.25	19.5	15.5
5	5	10	<1	1.59	10	<1	1.62	162	1.25	104.8	82.8
6	5	10	~1	2.55	10	<1	2.91	291	1.25	540.0	432
7	5	10	~1	0.40	10	<1	1.50	150	1.25	22.5	18
8	5	5	~1	2.90	5	<1	2.84	568	0.625	578.0	915
9	5	10	<1	0.56	10	<1	1.52	152	1.25	32.4	25.9
10	5	10	<1	1.10	10	<1	1.45	145	1.25	57.8	46
11	5	5	<1	0.62	5	~1	0.35	70	0.625	1.9	3

* Developed in channel resistance

** $G_p = \frac{P_o}{P_i} \simeq \frac{(\Delta I_D)^2}{(\Delta V_g)} \frac{R_c/8}{(\Delta I_g)/8}$

the channel dimensions as noted above. This, however, again introduces the problem of dimensional control.

The input power ΔP_i, consumed in the dc measurement, is totally dissipated in the shunt resistance of the gate electrode. This resistance arises almost solely from spurious leakage. Far smaller reverse currents are theoretically expected from a high bandgap material such as gallium arsenide.

The output power ΔP_o, as calculated in Table X, is completely absorbed in the channel resistance. While this is a somewhat artificial situation it gives a good indication of the signal power that could be produced in a matched load.

The power gain is computed from the simple "Class A" power amplifier relation shown at the bottom of table X. The gain values are indeed low as a result of the leaky input circuit. If V_{Do} is increased four-fold in Unit No. 1, as noted above, the gain value would be increased sixteen-fold making the gain 80. This, however, is also bought at the expense of high internal power dissipation.

The listed power gain values would apply up to the frequency at which the series resistance losses from capacitive displacement currents, in either the output or input circuits, become comparable with the shunt losses in the input circuit. This frequency would probably be in the hundred kilocycle region.

More detailed calculations on Unit No. 1 are compiled in Table XI. These are broadly illustrative of the performance to be expected from our sample units. The values are admittedly poor by today's bipolar transistor standards.

In Table XI the reason for discrepancy between the observed and

TABLE XI

GENERAL PERFORMANCE OF UNIPOLAR TRANSISTOR

Parameter	How Calculated	Value
1. Applied voltage (V_{Do})	-	0.85 volts§
2. Channel length (L)	-	0.05 cm§
3. Channel thickness (W_c)	R_c, dimens., and conduct	3×10^{-5} cm
4. Channel field (E_{co})	V_D/L	17 volt/cm
5. Carrier drift time (T_r)	$L/\mu E_c$	4.9×10^{-7} secs.
6. Current gain-bandwidth ($G_I \triangle f$)	$G_I \triangle f = \frac{1}{2\pi T_r}$	0.326 mes
7. Max. oscillation freq. (f_m)	$f_m = \frac{1}{2\pi R_o C_o}$	0.326 mes
8. Transconductance (g_m)	$g_m = \frac{\partial I}{\partial V}_{input} = \frac{Co}{2\,T_r}$	192 μmhos 50 μmhos§
9. Power gain-bandwidth $G_P (\triangle f)^2$	$G_P (\triangle f)^2 = \frac{1}{(2\pi T_r)^2}$	$(0.326)^2$ (mcs)^2
10. Modulation ($\triangle I_D$)	$\triangle I_D = I_{Do} \frac{W}{WC}$	10 ma 1 ma§
11. Active gate capacitance (C_o)	$C_o = 8.3 \times 10^{-5} \sqrt{\frac{EN}{V}} A$	188μμf
12. Depletion Layer Width W	$W = 1.05 \times 10^{3} \sqrt{\frac{EV}{N}}$	7×10^{-5} cm

§observed value

calculated values of the transconductance is probably that the channel impurity density is overestimated. Diffusion of the gate junction could well lead to partial compensation of the donors in the channel. This would decrease the gate capacitance and hence the transconductance. Furthermore, this error in channel impurity density would lead to a calculated value of channel thickness, W_c, that is too low. This, in turn, would lead to the sort of discrepancy between the observed and calculated values of ΔI_D noted in the table.

Improved performance could be obtained by reducing the channel length (increased transconductance) and by reducing the contact areas (reduced channel capacitance). By using current technology and materials we could probably improve the transconductance from 300 μmhos to 6000 to 10000 μmhos and improve the frequency response by a factor of more than a thousand.

J. Switching Transistor

1. Design Considerations

During the early part of this contact, a gallium arsenide crystal with satisfactory electrical and mechanical properties was not available. As soon as reasonable, single crystal gallium arsenide could be obtained, attempts were made to fabricate transistors using standard germanium and silicon techniques. The earliest attempts were made by alloying tin dots to opposite sides of a thin p-type gallium arsenide wafer that had been converted from n-type to p-type by copper diffusion. Some transistor action was observed but units were very poor. Around the same time, an attempt was made to make a diffused junction transistor. This was done by diffusing a p-type emitter and a p-type collector into opposite sides of an n-type gallium arsenide wafer. These units also showed some transistor action but were very poor. Typical current gains for these early units ranged from 0.01 to 0.05.

It was soon discovered that the minority carrier lifetime in gallium arsenide was too low for wide-based structures, and so work was started on an alloyed-emitter, diffused-base structure which could be made with base widths as small as 0.1 to 0.2 mils. Considerable effort was also put into making a double-diffused switching transistor with a base width of 0.015 to 0.03 mils. Unfortunately, a host of new problems were uncovered and the double-diffused switching transistor could not be made with high current gains.

Early attempts at making alloyed-emitter, diffused-base transistors were directed toward a simple mesa structure with an alloyed emitter dot and either one or two plated base contacts. These structures are shown in Figures 21 and 22.

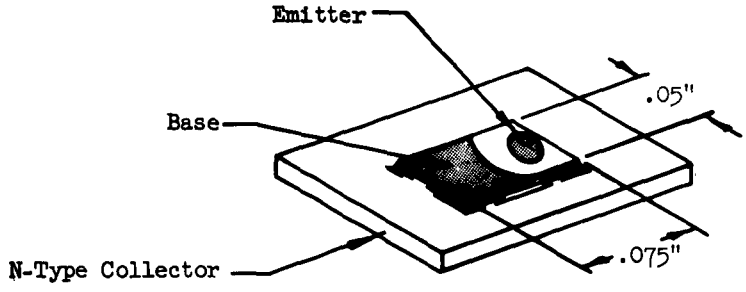


FIGURE 21 ALLOYED DOT EMITTER WITH SINGLE BASE CONTACT

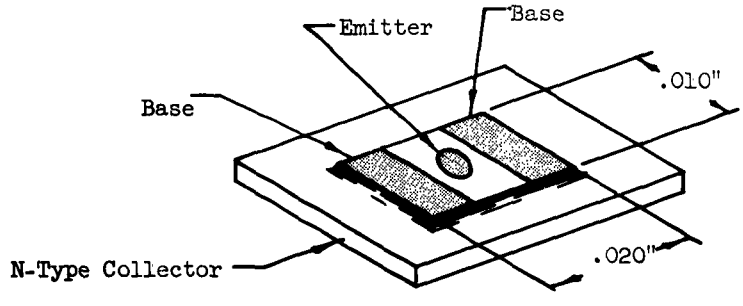


FIGURE 22 ALLOYED DOT EMITTER WITH DOUBLE BASE CONTACT

A theoretical comparison was made of sample germanium, silicon, and gallium arsenide structures to determine the potential of gallium arsenide high-frequency transistors. The results of this comparison shown in Table XII, prompted further investigation of the two stripe structure shown in Figure 23. In this structure, the two stripes were evaporated and alloyed, one as an emitter and one as a base. The concentration profile shown in Figure 24 was assumed where $W=X_e$. Although the values in Table XII look very optimistic, the structure could not be made to yield satisfactory characteristics. The two major problems were 1) the evaporated thin stripe would not alloy into a continuous emitter,

TABLE XII

COMPARATIVE CALCULATIONS ON Ge, Si, AND GaAs TRANSISTOR STRUCTURES

MATERIAL PROPERTY	GERMANIUM	SILICON	GALLIUM ARSENIDE	UNIT
Electron Mobility For 3×10^{16} /cm^3	2500	760	5000	cm^2/volt.sec.
Electron Mobility For 1×10^{17} /cm^3	1850	600	4200	cm^2/volt.sec.
Hole Mobility For 3×10^{16} /cm^3	1100	300	270	cm^2/volt.sec.
Hole Mobility For 1×10^{17} / cm^3	800	220	230	cm^2/volt.sec.
f_α for W = 0.04mils	1500	490	3400	mc
f_α for W = 0.1 mils	238	77	537	mc
f_α for W = 0.2 mils	60	19	134	mc
Effective lifetime in device for β=20 and W = 0.04 mils	2.1×10^{-9}	6.5×10^{-9}	9.4×10^{-10}	sec.
Effective lifetime in device for β=20 and W = 0.1 mils	1.3×10^{-8}	4.1×10^{-8}	5.9×10^{-9}	sec.
Effective lifetime in device for β=20 and W = 0.2 mils	5.4×10^{-8}	1.6×10^{-7}	2.4×10^{-8}	sec.
Measured effective lifetime	1×10^{-6}	5×10^{-7}	10^{-9}	sec.

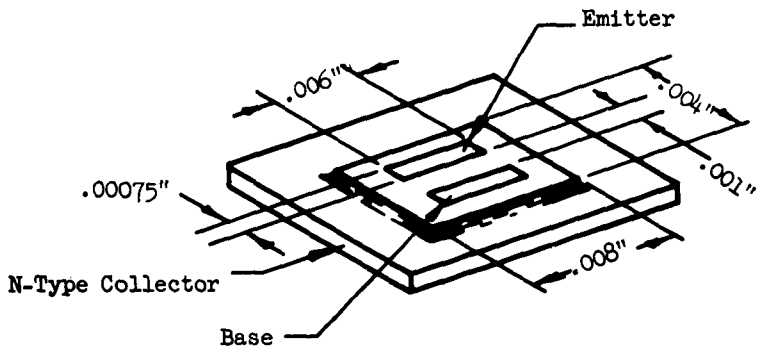


FIGURE 23 TRANSISTOR USING EVAPORATED EMITTER AND BASE CONTACTS

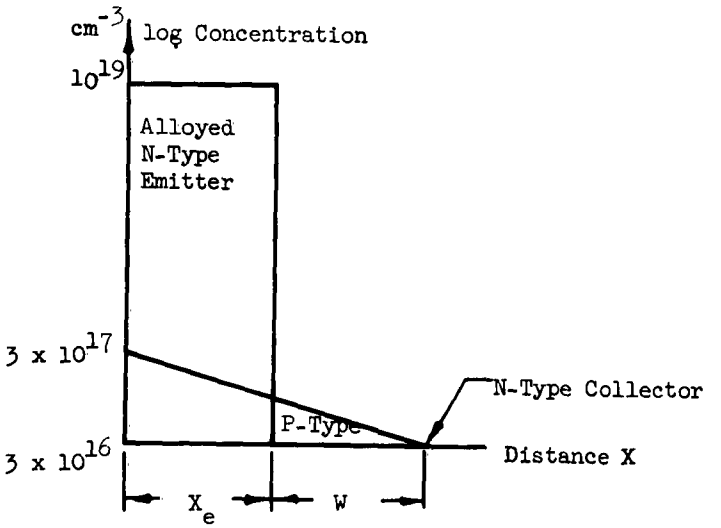


FIGURE 24 ASSUMED CONCENTRATION PROFILE FOR THE TWO-STRIP DEVICE

but would break up into many tiny metal dots, each alloying individually with the gallium arsenide and 2) thick stripes would not act as a satisfactory emitter. A maximum current gain of about 2 was achieved with this structure and even that at a very low yield level.

Even though this unit was not satisfactory, a set of tentative specifications were written for the gallium arsenide transistor showing

its intended application as a high-speed switching transistor. These specifications are shown in Table XIII.

Next, an alloyed-emitter, diffused-base mesa transistor with circular symmetry was considered. This structure appeared to have a number of advantages: 1) the circular emitter could be made by alloying a dot or by evaporating a heavy layer of metal and then alloying, 2) the ring base contact could easily be made by a variety of known processes and 3) the structure would be very suitable for high-frequency operation. The emitter injection efficiency for this structure can be calculated as:

$$Y = \frac{1}{1 + \frac{\rho_e W}{\rho_b L_{pe}}} \qquad (10)$$

Where: ρ_b (base resistivity) $\cong R_s \, x_o$ Ω - cm

L_{pe} (diffusion length for holes in emitter region)

$$= \sqrt{D_p \, \tau_p} = \sqrt{\frac{k\,T}{q} \, \mu_p \, \tau} \quad \text{cm}$$

ρ_e (emitter resistivity) $= \frac{1}{q \, n \, \mu_n}$ Ω - cm

W (base thickness) = 5×10^{-4} cm = 0.2 mils

For a typical transistor:

R_s (sheet resistivity of base) $\approx 500 \Omega/\square$

x_o (diffusion depth) $\approx 7.5 \times 10^{-4}$ cm

μ_p (mobility for holes) ≈ 100 cm^2/volt-sec at 10^{19}/cm^3 doping level.

τ (lifetime of holes) $\approx 10^{-10}$ sec

n (concentration of carriers in emitter) $\approx 10^{19}$

μ_n (mobility for electrons) ≈ 1000 cm^2/volt-sec

TABLE XIII

TENTATIVE SPECIFICATIONS FOR GaAs SWITCHING TRANSISTOR

BV_{CBO}		15 - 40 volts
I_{CO}		5 na
h_{fe}		18
C_{ob}		3 pf
R_{CS}		20 ohms
Max Junction Temperature		300°C
Turn on time	t_d	5 ns
	t_r	13 ns
Turn off time	t_s	8 ns
	t_f	24 ns

Substituting these approximate values, the emitter efficiency is 0.951.

Thus, the alloyed tin emitter would be satisfactory if it were not its temperature limitations (tin melts at 232°C). The temperature problem was solved by using a gold-tin (50-50 atomic percent) dot which melts at 418°C.

A number of minor modifications were made on this basic structure such as variations in dimensions, in the technique of applying the base contact, and in the technique of forming and alloying the emitter, but the basic structure is still very similar to the original. The final structure adopted for the switching transistor is shown in detail in Figure 25. Both the emitter and base, in the final structure, are made by an electroplating process followed by a heating step which alloys the emitter and sinters in the base contact. A more detailed description of the processing techniques, problems, and characteristics of this structure will be given under device processing.

To reduce the base width of the alloyed-emitter, diffused-base transistor, the post alloy diffused (PADT) approach was tried. The physical structure of the PADT transistor is the same as that of the alloyed-emitter, diffused-base transistor and much of the processing is the same. The two major differences are: 1) the diffused p-layer is made very thin so that the emitter will alloy through it and 2) the emitter contains a p-type diffusant in addition to an n-type material to dope the regrowth. In assembling this transistor, the double-doped emitter dot is alloyed at a high temperature so that a p-type base diffuses just in front of the interface between the emitter dot and the solid gallium arsenide. On cooling, the n-type regrowth forms the emitter, the very thin p-type

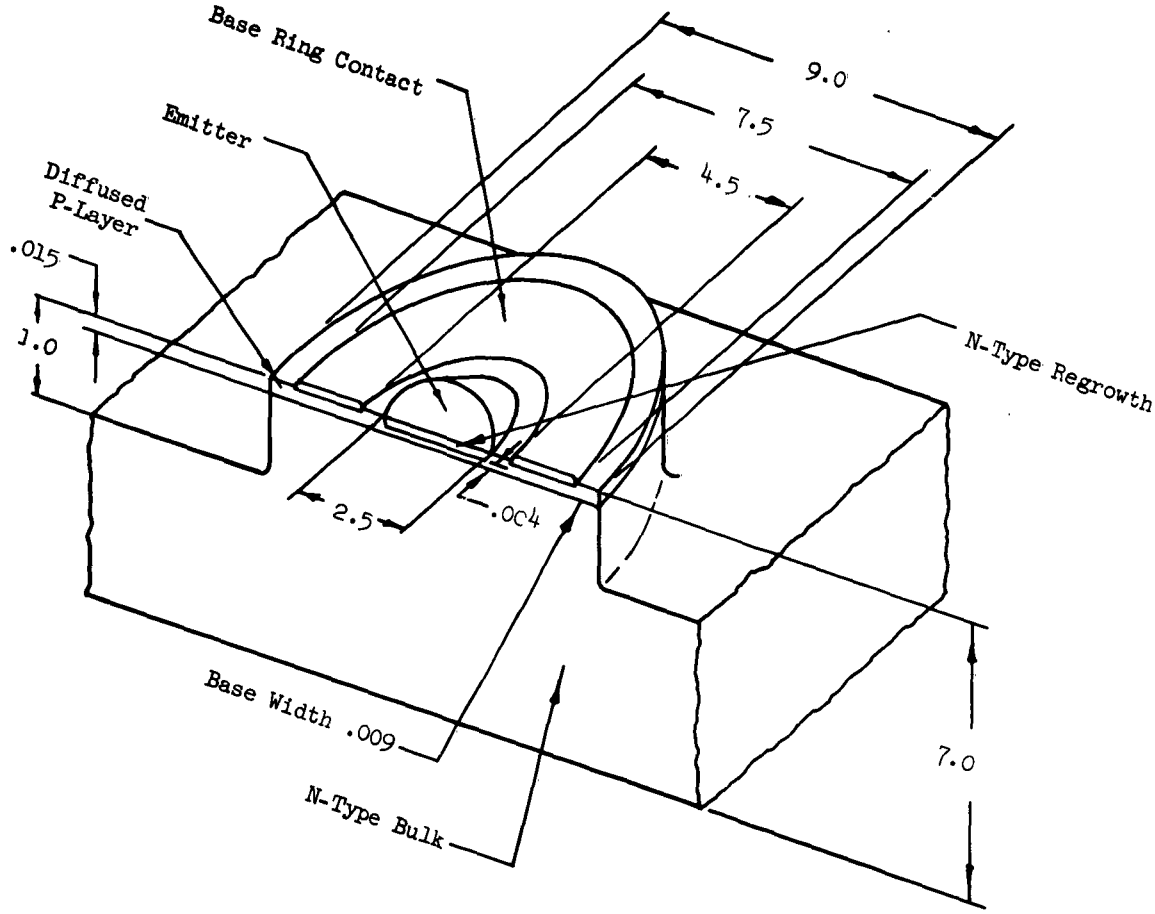


FIGURE 25 ALLOYED-EMITTER, DIFFUSED-BASE SWITCHING TRANSISTOR, FINAL STRUCTURE

diffused layer forms the base and the junction between the diffused region and the bulk gallium arsenide forms the collector. The major advantage of this type structure is the ability to form extremely thin and uniform base regions.

The PADT approach works quite well with germanium and there was every reason to believe it would also work with gallium arsenide. An extensive effort was put into developing this structure, but satisfactory

transistors could not be made. The reasons were not known at the time but based on later data and a better understanding of the behavior of gallium arsenide and the various elements which dope it, a satisfactory explanation can now be given. The n-type materials used were tin and gold-tin (50-50 atomic percent) and the p-type materials were zinc, cadmium, and manganese with all six possible combinations tried. It was found that if there was a low percentage of p-type dopant in the dot, the dopant would not diffuse out to form a p-type base region. On the other hand, large amounts of dopant would diffuse out, but the resulting transistors either were very poor or in many cases would not work at all.

Further investigation of this process yielded the following facts: 1) the solubilities of manganese and cadmium in gallium arsenide are as high as the solubility of tin, while the solubility of zinc is an order of magnitude higher, 2) the diffusion coefficients of zinc and cadmium are both anomolous and show a very strong dependence on concentration and 3) manganese reacts with tin in gallium arsenide crystal and ties up both the tin and manganese resulting in a compensated emitter and a very slow diffusion of manganese from the emitter.

The explanation for the poor performance of the PADT transistor is thus believed to be: 1) when manganese was used it reacted with the tin in the emitter and therefore, was not able to diffuse out and form a base region; 2) when zinc was used, it either over doped the emitter making it p-type or if a low concentration was used, it would not diffuse because of the very low diffusion coefficient for zinc at low concentrations; 3) when cadmium was used, the situation was essentially the same as for zinc.

With an npn structure using a base width of 0.03 mils, a base doping of

10^{17} carriers/cm^3, base lifetime of 10^{-8} seconds, emitter doping, Co, of 2 x 10^{19} carriers/cm^3 and emitter lifetime of 10^{-10} seconds, the emitter injection efficiency limited current gain is 20. For a pnp structure with a base width of 0.01 mils, a base doping of 5 x 10^{17} carrier/cm^3, base lifetime of 2 x 10^{-9} seconds, an emitter doping Co, of 10^{20} carrier/cm^3 and an emitter lifetime of 10^{-11} seconds, the emitter injection efficiency limited current gain is about 90.

These lifetime values are now believed to be quite optimistic, but the true values still cannot be measured. It is probable that some of the assumed values are at least an order of magnitude too high, based on electrical and mechanical measurements of finished transistors. The npn structure was made as shown in Figures 26 and 27. Figure 26

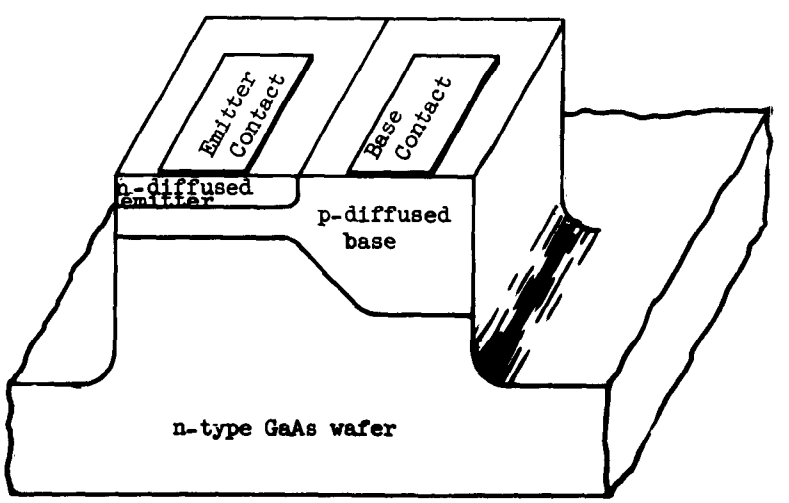


FIGURE 26 DOUBLE-DIFFUSED SWITCHING TRANSISTOR WITH STRIPE CONTACTS

shows a linear geometry with stripe contacts while Figure 27 shows a circular geometry essentially like that used in the alloyed-emitter

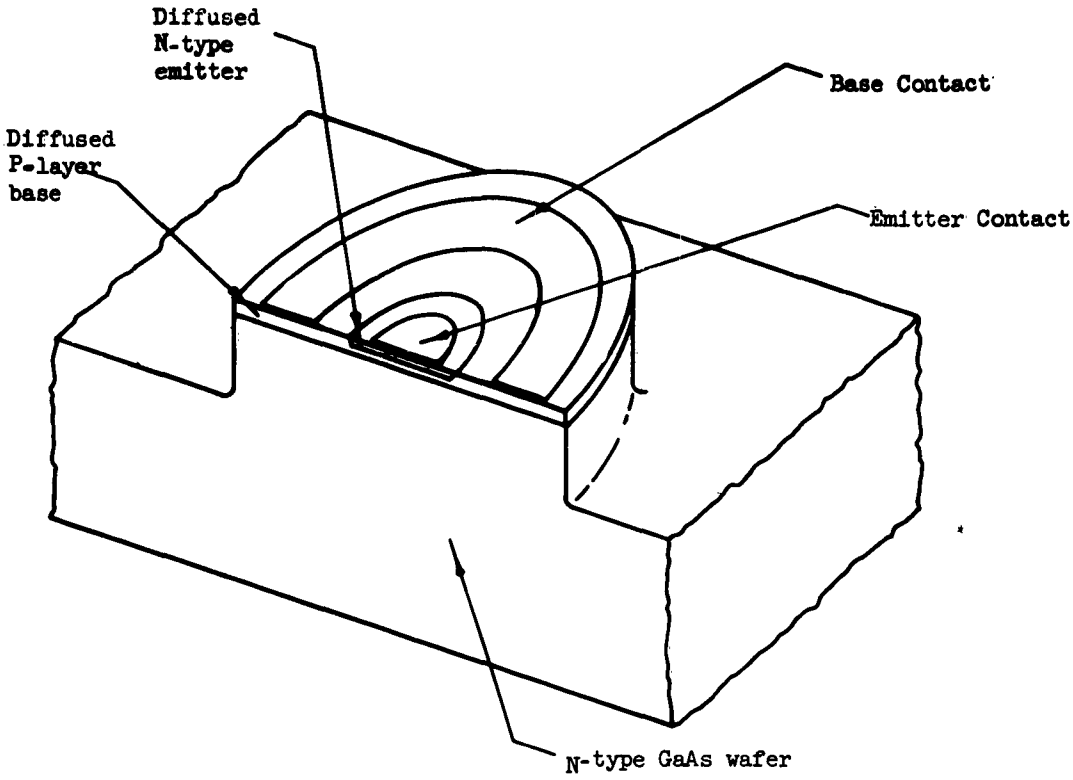


FIGURE 27 DOUBLE-DIFFUSED SWITCHING TRANSISTOR WITH CIRCULAR GEOMETRY

diffused base structure. The linear geometry of Figure 26 was used with tin and manganese as diffusants while the structure of Figure 27 was used with tin and zinc as diffusants. The tin-manganese units exhibited a maximum current gain of about 7 while the tin-zinc units exhibited a maximum current gain of about 4. Electrical measurements on these units indicated that they were limited by emitter injection efficiency. This is consistent with later evidence that the assumed minority carrier lifetime values used in the calculation are too high by at least an order of magnitude. The only way to improve the double-diffused transistor would be to improve the minority carrier lifetime, increase the doping ratio between emitter and base or reduce the base width. Since none of these could be done to any significant degree,

the double-diffused transistor cannot be made with the presently available materials and techniques.§

All work on the pnp double-diffused transistor was done on the power transistor structure and will be reported there.

After a careful consideration of all possible structures, the alloyed-emitter, diffused-base transistor was chosen as the structure most likely to result in useful transistors. All samples submitted were made with this structure.

2. Device Processing

For the alloyed emitter structure, the starting crystal must be single and oriented on the (111) face with a carrier concentration of between 1 x 10^{16} and 8 x 10^{16} carriers /cm^3 and a dislocation density of from 5,000 to 100,000 per cm^2. The major problems with gallium arsenide crystal are the very low minority carrier lifetime and the very poor crystal uniformity. True minority carrier lifetime is so low that it has never been measured without very large errors and ambiguties. In addition, the normal properties of gallium arsenide such as resistivity, carrier concentration, dislocation density and diffusion coefficients for impurities vary so much from crystal to crystal and from point to point within a given crystal that each crystal must be individually studied to determine its properties.

It has been found, in general, that vertically grown crystals result in somewhat better transistors even though the lifetime, dislocation density, and mobility seem to be better on horizontally grown crystals. This anomaly has never been solved. In addition, extensive experiments have failed to show any significant correlation between short order uniformity in crystals and properties of

§If wide band gap emitters in gallium arsenide become a reality good all diffused or epitaxial transistors will be made.

transistors made from these crystals.

Of all processing areas involved in making a transistor, diffusion was probably the most thoroughly studied. A total of nine diffusants were studied in considerable detail and at least seven more were investigated to a lesser extent. Table XIV lists the diffusants studied in detail together with some of the significant properties of each. Of the p-type diffusants, zinc is most suitable for high concentrations and manganese for low concentrations. Of the n-type diffusants, tin is most suitable for high concentrations and either tellurium or sulfur for low concentrations. Figure 28 shows diffusion coefficients for several diffusants or a function of temperature and Figure 29 shows the anomalous behavior of zinc at 900°C. For the alloyed-emitter, diffused-base transistor, manganese is evaporated onto the reverse side of the gallium arsenide and the diffusion is done in a sealed quartz ampoule under high vacuum at about 750°C. Diffusion time is varied for each crystal but runs typically from about 6 minutes for fast diffusing crystals to about 40 minutes for slow diffusing crystals. Typical junction depths are 0.15 mils and typical sheet resistivities are 700 ohms per square. Epitaxially grown gallium arsenide layers usually diffuse very slowly compared to conventional crystal.

In making alloyed-emitter, diffused-base transistors, the next step after base diffusion is the preparation of the base and collector contacts. For the n-type collector contact, nickel is plated over an evaporated tin layer and the contact is then sintered in hydrogen at 600°C. For the p-type base contact, early units had gold--1%

TABLE XIV

DIFFUSANTS FOR GaAs

Diffusant	Type	Useful Concentration	Properties
Zn	p	10^{18} to 10^{20}	Diffusion coefficient is a very strong function of concentration.
Cd	p	10^{18} to 10^{19}	Diffusion coefficient is a function of concentration. Severe surface attack on GaAs.
Mn	p	10^{17} to 10^{18}	High ionization energy. Diffusion fronts often very rough.
Se	n	5×10^{18}	Glass formation and surface attack on GaAs. Self compensating at less than 10^{19} carrier/cm^3.
Te	n	5×10^{18}	
S	n	10^{18}	Surface attack on GaAs
Si	n	10^{17} to 10^{19}	Surface attack on GaAs Self compensating
Ge	n	5×10^{18}	Self compensating. Results in low mobility in crystal
Sn	n	10^{18} to 2×10^{19}	Well behaved error function diffusion.

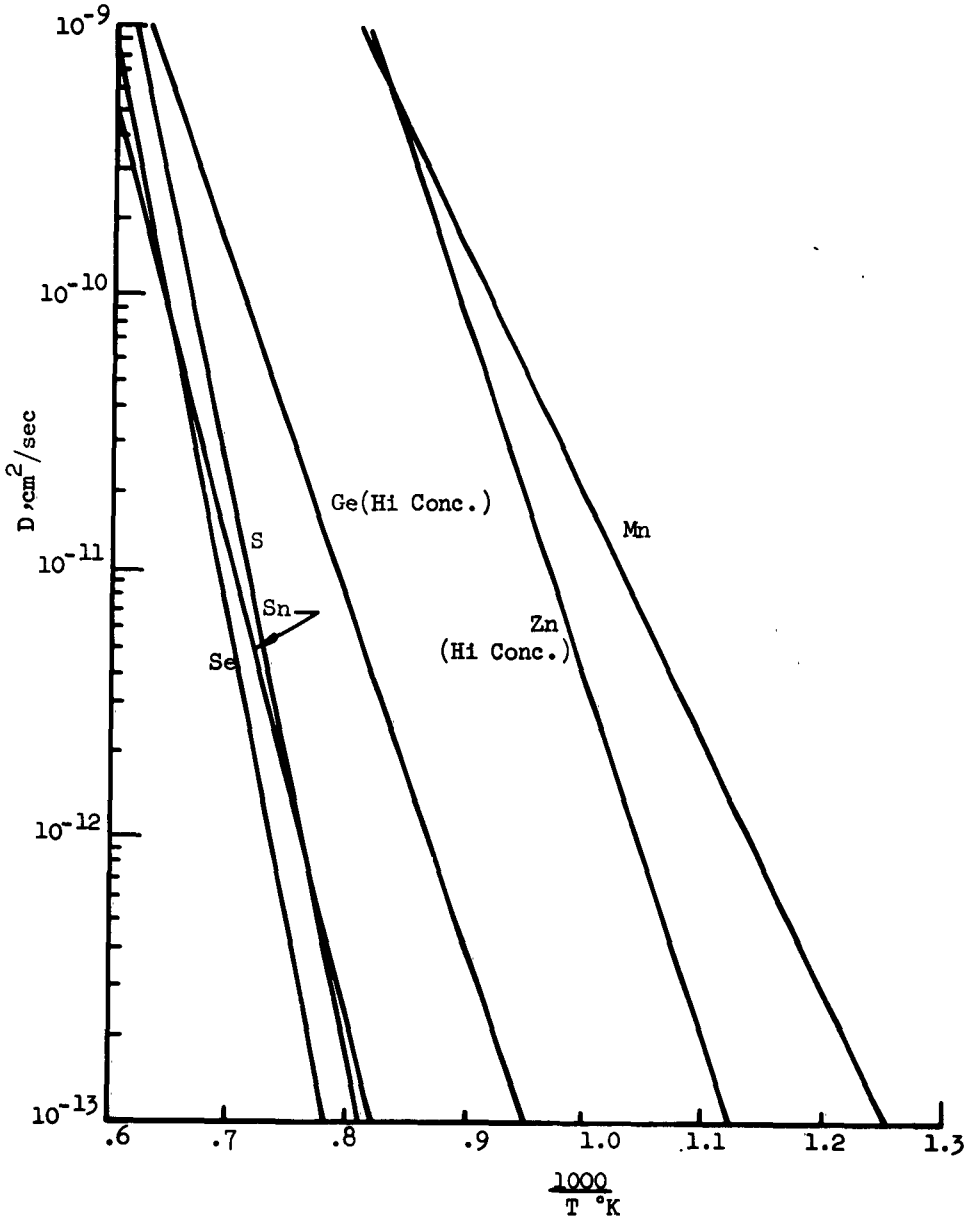


FIGURE 28 SOME DIFFUSION COEFFICIENTS FOR GALLIUM ARSENIDE

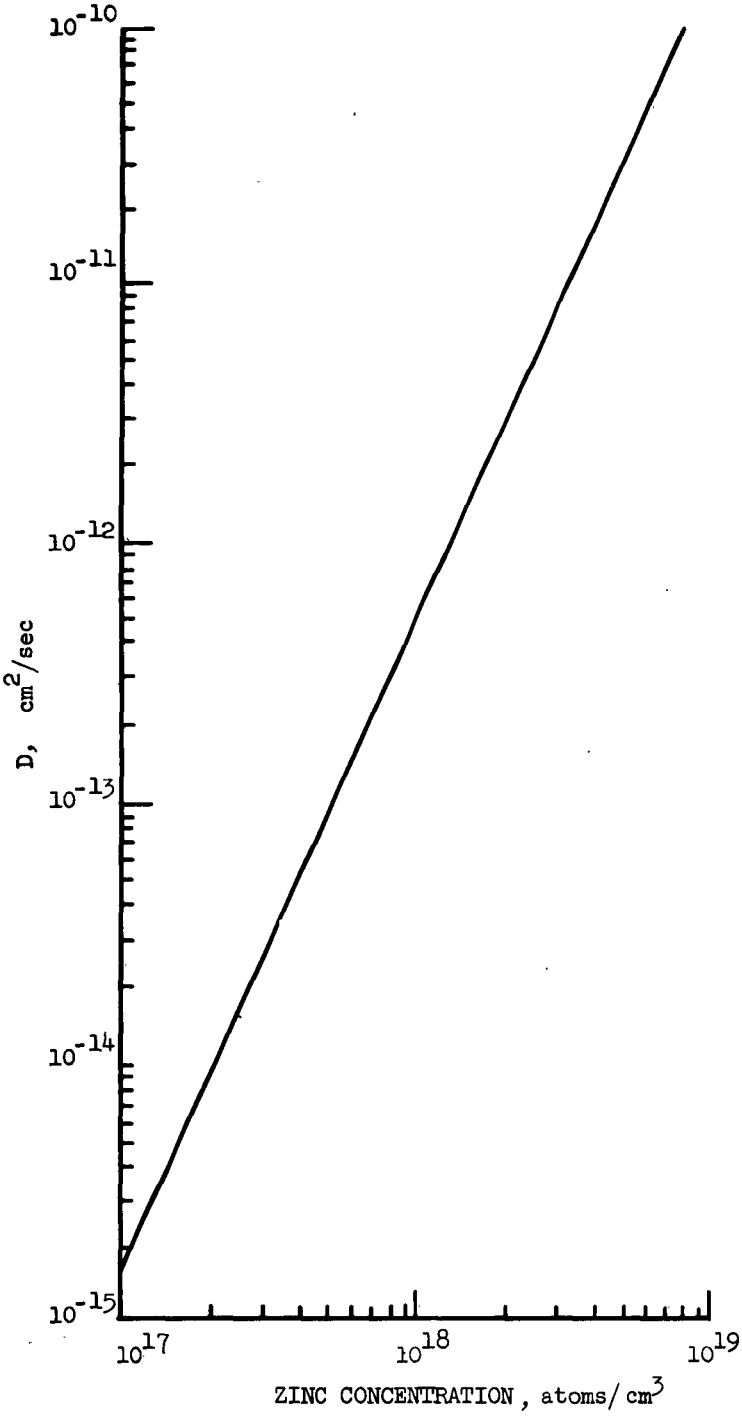


FIGURE 29 DIFFUSION COEFFICIENT FOR ZINC AS A FUNCTION OF ZINC CONCENTRATION (900°C)

manganese evaporated over a nickel plating. Present processing calls for electroplated pure silver. In either case, the base contact is also sintered in hydrogen at 600°C, usually at the same time as the collector contact. The base contact is defined by coating the gallium arsenide wafer with silicon dioxide and using photolithographic techniques to remove regions of the oxide where a contact is desired. A number of other n-type and p-type contacts were investigated but none were both as convenient to apply and as good electrically as the ones chosen.

The next major operation after formation of the base and collector contacts is the formation of the emitter. Originally, emitters were put on in the form of small (2 mil) dots and alloyed in a furnace containing a hydrogen atmosphere. Table XV shows a list of various emitter alloys which were evaluated. The gold-tin compound melting at 418°C was chosen as the best in terms of electrical properties, melting point and ease of making and handling. After gold-tin was selected as the emitter material, it was studied thoroughly not only as a dot material but as an evaporated material and as an electroplated material. The properties of the material are the same regardless of the technique of application if a sufficiently large quantity of alloy is used. Evaporated and plated dots are not found to be satisfactory with thin layers.

A major problem was encountered in controlling the penetration and spreading ratio of gold-tin dots. This problem was partially solved by 1) using electroplated material with a silicon dioxide mask on the surface of the gallium arsenide to prevent or at least limit spreading and 2) alloying in an atmosphere containing excess arsenic vapor to limit the decomposition of the gallium arsenide. The emitter on the alloyed-

TABLE XV

METALLURGICAL AND ELECTRICAL PROPERTIES OF EMITTER ALLOYS

ALLOY	MELTING POINT °C	COMMENTS
AuTe	447	Rectifying on p-type, no regrowth
TePb	405	Rectifying on p-type, no regrowth
AuInSb	451	Rectifying on p-type, no regrowth
PbPdSn	450	Ohmic on p-type
SbInTe	535	Rectifying on p-type
SbInSe	535	Rectifying on p-type
AuGeSn	356	Rectifying on p-type
SbTe	540	
PbPdSe	450	
SnAu	418	
PbPdTe	450	
AuTeSe	447	
AuInSe	451	
AuInTe	451	
TeBiSe	413	
TeIn	667	Poor wetting
TeBi	413	
AuPbTe	451	
AuSnZn	418	
AuSnMn	418	
AuSnS	418	
AuSnTe	418	Difficult to shoot
AuSnSe	418	Difficult to shoot
SnSe	232	
SnTe	790	Poor wetting
PbSe	1088	
SbSn	-	Poor wetting
GeSn	-	
SbAuGe	-	Germanium regrowth

emitter, diffused-base transistor is made by electroplating gold and then tin through a silicon dioxide mask and then alloying in hydrogen containing a few millimeters arsenic pressure. The arsenic pressure can be maintained by using either pure arsenic with a two temperature zone furnace or using a compound such as germanium arsenide with a suitable arsenic vapor pressure at the alloying temperature.

The electrical properties of alloyed emitters have never been as good as those of diffused emitters and despite a considerable effort in this area, the problem was not completely solved. It is probable that a number of factors contribute to the problem. The difficulty in obtaining good regrowth under the emitter is certainly important. Often there is a component of temperature independent excess current,[§] not representing carrier injection, and a component of leakage across the junction. It is very probable, however, that some additional factor is involved and it may be some sort of a reaction between the manganese in the base and the tin in the emitter. Junctions between tin and zinc are generally better than junctions between tin and manganese.

After emitter alloying, mesas are formed using standard evaporated wax masking techniques and etching. Pellet scribing, mounting and bonding are all standard techniques and are done in the same way for gallium arsenide as for germanium or silicon transistors. The pellet solder for attaching the pellet to the TO-18 header is a ternary alloy of gold, indium, and tin melting at 451°C. The emitter lead connection is 0.5-mil platinum wire and the base connection is 0.5-mil gold wire. The mounting and bonding are done on Kulicke and Soffa machines.

After mounting and bonding, the collector junction often shows

[§]Such as in gallium arsenide tunnel diodes

an increased leakage current. To reduce the leakage current and also to improve current gain, a clean up etch is included at this point. A wide variety of etches, both acid and electrolytic were investigated. In general, it is possible to restore low leakage, to a unit, but it generally is not possible to effect any permanent improvement in current gain. The etch which was found to yield the best result is an electrolytic etch in potassium hydroxide solution. In some cases, the diffusant contact materials act as electrolytic cells so that no external bias is necessary. Acid etching is also effective in improving collector junctions, but, this type of etching is not used because the etch attacks the metal header and metallized contacts.

After etching, gallium arsenide switching transistors are baked in air at temperatures ranging between 120°C and 200°C and then sealed with dry nitrogen. There seems to be little significance in the temperature or atmosphere used for the pre-seal baking.

Final testing is performed on standard commercial test sets. Test equipment and techniques are no different than for silicon or germanium transistors and thus need not be described in detail.

The characteristics of samples submitted under this contract are shown in Table XVI.

TABLE XVI

CHARACTERISTICS OF NPN GALLIUM ARSENIDE SWITCHING TRANSISTOR SAMPLES

Unit No.	BV_{CBO}, volts	I_{CO}, μa	h_{fe},	C_{ob}, pf	V_{CE}, volts	R_{CS}, ohms	Switching Times t_d ns,	t_r ns,	t_s ns,	t_f ns,
1	18	.0008	29.5	8.8	-	45	16	24	8	16
2	16.5	.064	20.2	6.7	-	18	12	20	4	16
3	30	.0097	13.3	4.8	-	45	12	28	4	12
4	48	.0004	11.8	8.4	-	38	12	20	4	12
5	27	.0004	11.0	8.7	-	42	16	16	4	12
6	11	.028	18	3.0	-	80	32	48	12	20
7	11	.10	42	6.5	-	30	36	24	16	20
8	16	.044	17	5.5	0.8	-	12	32	4	28
45	12	.03	19	3.8	0.45	-	3	20	9	15
43	37	.0085	11	4.8	0.6	-	4	20	7	20
16	3	-	37	-	0.3	-	4	16	20	40
31	7	26.	18	4.0	0.7	-	8	16	12	30
10	25	.12	12.5	4.8	1.2	-	8	20	8	18
32	18	.054	11	4.2	1.2	-	8	24	10	26
33	10	.013	12	2.8	1.0	-	8	28	8	24
42	60	.0003	8	5.2	1.1	-	4	27	8	20
12	30	.013	9	5.2	0.6	-	8	24	8	20
18	37	.0002	4.5	6.3	0.75	-	4	30	4	20
15	13	.018	8.5	3.0	1.25	-	8	25	6	18

BV_{CBO} measured at $I_C = 100\mu a$

I_{CO} measured at $V_{CB} = 5$ volts

h_{fe} measured at $V_{CE} = 5$ volts, $I_E = 10ma$, f = 1 Kc

C_{ob} measured at $V_{CB} = 5$ volts

V_{CE} measured at $I_C = 10ma$

Switching times are measured in a circuit with a forced beta of two. The rise time of the circuit is 12 ns and these values must be corrected for this error. The formula for this correction is: $t_r = \sqrt{t_{tot}^2 - t_{eq}^2}$

K. Power Transistor

1. General

As the state-of-the-art in gallium arsenide advanced to the point where dopant diffusions and photolitographic techniques could be applied successfully to this material, the development of a power transistor was considered. It was clearly seen, that only a double-diffused structure would yield useful devices, since a large area emitter junction with a high ratio of periphery length to contact area (either long narrow stripe or comb structure) cannot easily be formed by evaporation and alloying. The high temperature alloy (gold-tin) stripes tend to "ball up" resulting in discontinuities and uneven regrowth. Furthermore, the close base width control ($W \leq 10^{-4}$ cm), which is necessary to meet the high-frequency requirements of the power transistor, can be achieved reproducibly only by diffusion.

The pnp-approach was investigated first, since it was observed in previous work that in gallium arsenide donors generally have a smaller diffusion constant and lower solubilities than acceptors. This suggests a slow n-type base diffusion with moderate surface concentration and a fast p-type emitter diffusion with high surface concentration. Design calculations (see Section 2a) revealed that an adequate emitter efficiency was obtainable despite the unfavorable ratio of the hole-to-electron mobility.

A number of pnp-units were fabricated. The current gain of these samples was less than unity, although good emitter and collector characteristics were achieved. In addition, the geometry control was so refined that a base width of $W = 0.6 \times 10^{-4}$ cm was obtained consistently. Analysis led to the conclusion that the minority carrier lifetime was several orders of magnitude shorter than anticipated from the theoretical limit of band-to-band recombination ($\tau_{eff} < 10^{-11}$ sec; $\tau_{theoretical} \sim 10^{-8}$ sec).

To circumvent the above problem, a low concentration donor diffusion process was developed to utilize the fact that for zinc the diffusion constant decreases considerably as concentration is lowered. This provdied the possibility of building an npn-structure. Although the optimum concentration profile for the device could not be achieved, because of material and processing limitations, samples exhibited current gains from 2 to 7 and power gains in excess of 20db at f = 50 mc. A maximum power output of about 400 mw in Class A operation was measured.

It is evident from measurements and device analysis that the performance of the power transistor can be further improved by: a) Changing the geometry from the three-stripe configuration to a more elaborate interdigited one. This will lead to a better ratio of emitter periphery to emitter area. b) Using epitaxial material, because this is not as conversion sensitive at low carrier concentrations. Diffusion experiments with material having a carrier concentration of about 4×10^{15} carriers/cm^3 have shown breakdown voltages in excess of 100 volts. c) Planar base diffusion techniques have the potential of lowering collector leakage current and preventing early surface breakdowns. d) The use of a wide band gap emitter (i.e. gallium-phosphide) should vastly increase the emitter efficiency.

2. Design Considerations

The specification for the gallium arsenide power transistor is given in Table XVII.

Experience gained from silicon devices with similar requirements helped to select the device geometry.

The geometrical configuration and physical dimensions of the unit are shown in Figure 30. It is a double-diffused structure having a diffused emitter, 2 x 18 mils, with two base contacts, 1 x 18 mils. The contacts are spaced 1 mil from the emitter on either side. The mesa area is 8 x 20

TABLE XVII

DEVICE SPECIFICATION FOR GALLIUM ARSENIDE POWER TRANSISTOR

Collector Breakdown Voltage (I_C = 1 ma)	40 volts
High Frequency Current Gain (f = 100 mc)	2
High Frequency Power Gain (f = 50 mc)	11 db
Maximum Output Power (f = 50 mc, T = 200°C)	1 watt
Collector Capacitance ($V_{CB} \geq$ 10 volts)	8 pf
Maximum Operating Temperature	300°C

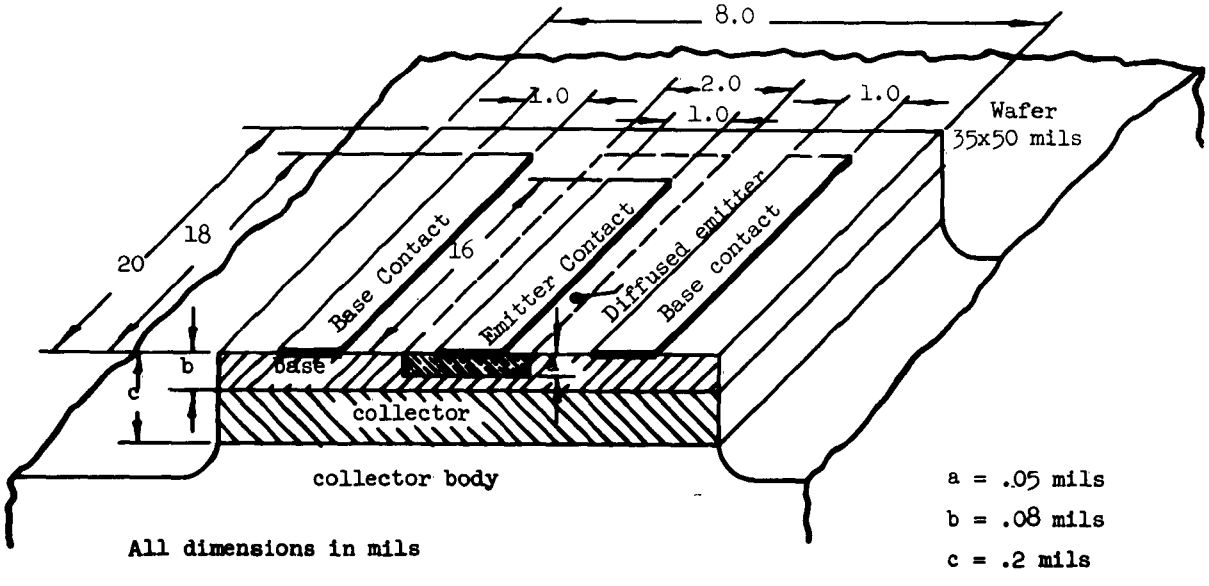


FIGURE 30 GALLIUM ARSENIDE POWER TRANSISTOR

mils; the gallium arsenide pellet is 35 x 50 mils.

a. PNP-Approach

(1) Impurity Profile

The net impurity density for the double-diffused structure (assuming the diffusion process obeys an error function distribution)

can be described by equation (11).

$$N(x) = N_{10}\,\mathrm{erfc}\left(\frac{x}{L_1}\right) - N_{20}\,\mathrm{erfc}\left(\frac{x}{L_2}\right) + N_3 \qquad (11)$$

There are three important factors which restrict the impurity profile. First, the emitter must be largely compensated to obtain high emitter efficiency. Therefore, a heavily doped emitter region and a lightly doped base layer are required. Second, for a high reach-through voltage from collector-to-emitter a minimum net-doping is necessary in the base, thus opposing the first requirement. Third, the surface concentration for the base and emitter diffusion must be within the limit of solid solubility. For these reasons the following parameters were selected for equation (11):

$N_{10} = 1 \times 10^{20}$ carriers / cm^3, zinc surface concentration

$N_{20} = 8 \times 10^{17}$ carriers / cm^3, sulfur surface concentration

$N_3 = 5 \times 10^{16}$ carriers / cm^3, cadmium doping level of p-type starting material

$x_{je} = 1.2 \times 10^{-4}$ cm, distance of emitter junction from surface

$x_{je} = 2 \times 10^{-4}$ cm, distance of collector junction from surface

$W_o = x_{jc} - x_{je} = 0.8 \times 10^{-4}$ cm, physical base width

The base diffusion length is determined by equation (12):

$$L_2 = \left(D_{T_2}\ t_2\right)^{1/2} \qquad (12)$$

where:

L_2 = diffusion length

D_{T_2} = diffusion constant donor atoms at temperature T_2

t_2 = diffusion time

Using the given values, the diffusion length

$$L_2 = 1.52 \times 10^{-4} \text{ cm.}$$

The emitter diffusion length is found in a similar fashion to be

$$L_1 = 0.538 \times 10^{-4} \text{ cm.}$$

Now all quantities in equation 11 are known and the profile can be computed. In Figure 31 the absolute value of the net density N (x) is shown as a function of the distance from the surface.

(2) Emitter Efficiency

As pointed out previously, the emitter efficiency is of vital importance. Since it is one of the factors limiting the current gain of the device, it should be as close to unity as possible. The emitter efficiency can be defined by:

$$\gamma = \frac{1}{1 + \frac{D_n}{D_p} \frac{J_1}{J_2}} \tag{13}$$

where:

D_n = Diffusion constant for electrons

D_p = Diffusion constant for holes

$$J_1 = \int_{x_{je}}^{x_{jc}} N(x)\, dx, \tag{14}$$

$$J_2 = \int_{0}^{x_{je}} N(x)\, dx, \tag{15}$$

Integrating equations (14) and (15) one finds that the numerical values are:

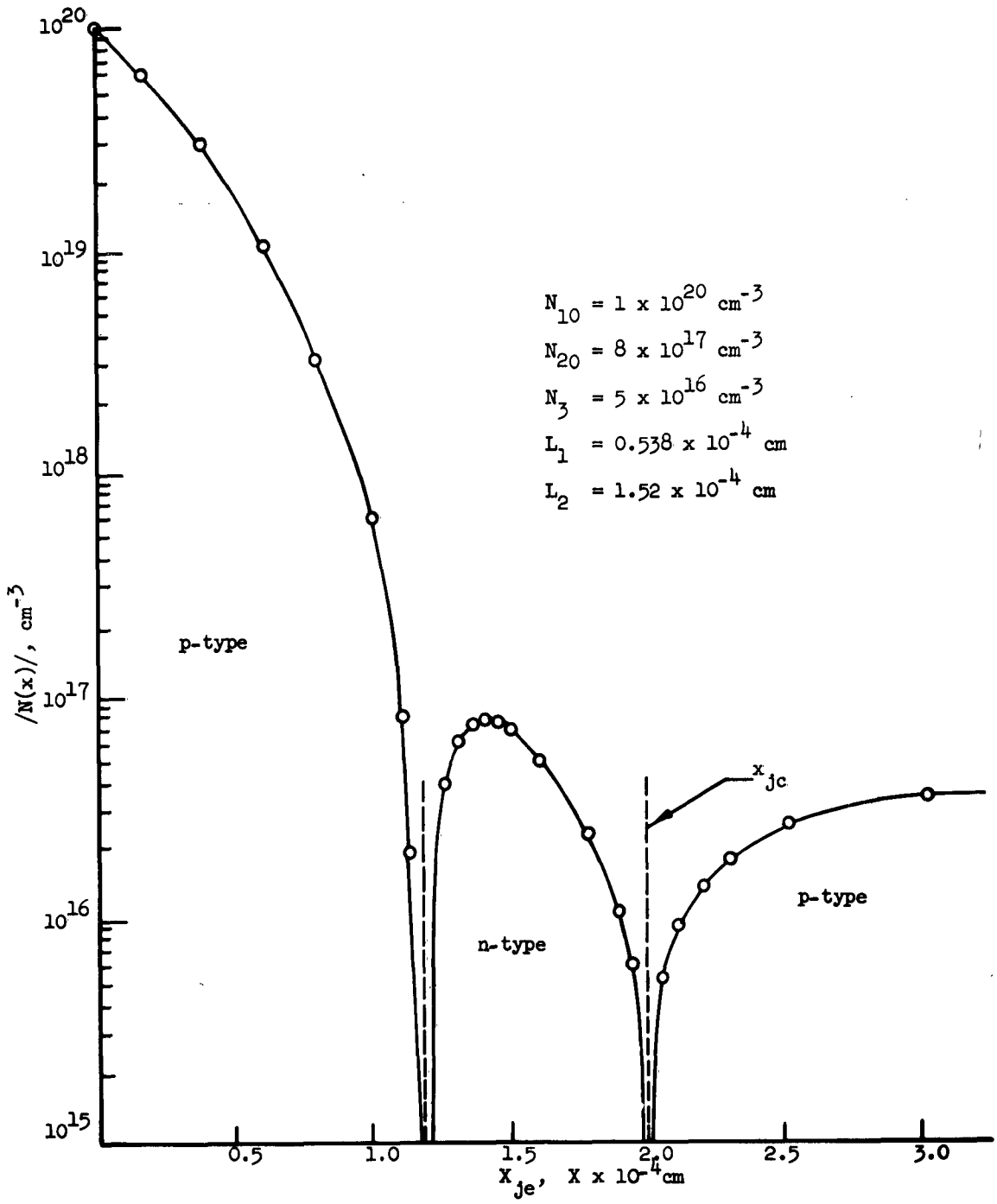


FIGURE 31 NET IMPURITY DENSITY FOR GaAs-POWER TRANSISTOR $/N/=f(x)$

$$J_1 = 3.381 \times 10^{12} \text{ cm}^{-2}$$

$$J_2 = 2.96 \times 10^{15} \text{cm}^{-2}$$

The emitter efficiency can be calculated from equation (13), with D_N = 125 $\text{cm}^2 \text{ sec}^{-1}$ and D_p = 9.6 $\text{cm}^2 \text{ sec}^{-1}$.

$$\gamma = \frac{1}{1 + \frac{125}{9.6} \times \frac{3.385 \times 10^{12}}{29.62 \times 10^{14}}} = 0.987 \qquad (16)$$

Therefore, the γ limited current gain (grounded emitter) becomes

$$\beta = \frac{\gamma}{1 - \gamma} = \frac{0.987}{0.013} = 76 \qquad (17)$$

As can be seen from the results of these calculations, a gallium arsenide transistor having a pnp structure is feasible under the above conditions.

(3) Collector Region and Base Layer

The design parameters for the collector region and base layers can be calculated using normalized curves as shown in Figure 32.(ref. 12) The numerical values for the normalized quantaties are in this case:

$$E_N = 4.7 \times 10^5 \text{ volts/cm}$$

$$V_N = 27.1 \text{ volts}$$

$$C_N = 17 \times 10^{-9} \text{ farads cm}^{-2}$$

Figure 33 shows the calculated depletion layer extending into the base (W_1), into the collector (W_2), the effective base width (W_{eff}) and the collector capacitance (Cob) as a function of voltage. For a bias voltage of V_C = 20 volts we get:

$$\text{Cob} = 8.9 \text{ pf}$$

$$W_{eff} = 0.29 \times 10^{-4} \text{cm}$$

At V_C = 40 volts, the depletion layer has reached the region of maximum

ref. 12 R.M. Scarlett, "Space-Charge Layer Width in Diffused Junctions", IRE - PGED, Transaction 1959

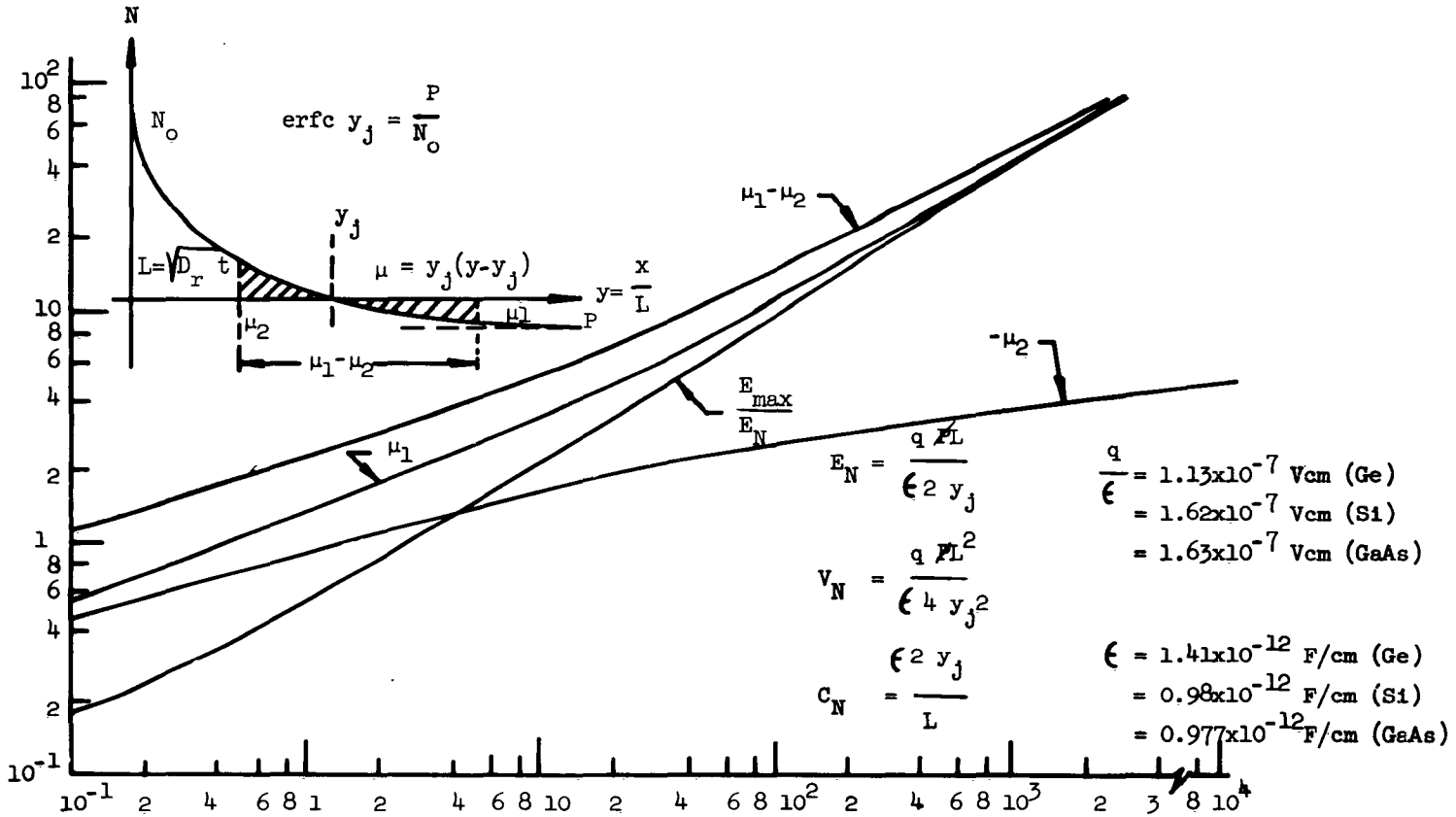


FIGURE 32

SPACE-CHARGE LAYER WIDTH, MAXIMUM FIELD, AND JUNCTION CAPACITANCE VS. JUNCTION VOLTAGE FOR APPROXIMATION TO COMPLEMENTARY ERROR FUNCTION (ALL QUANTITIES NORMALIZED)

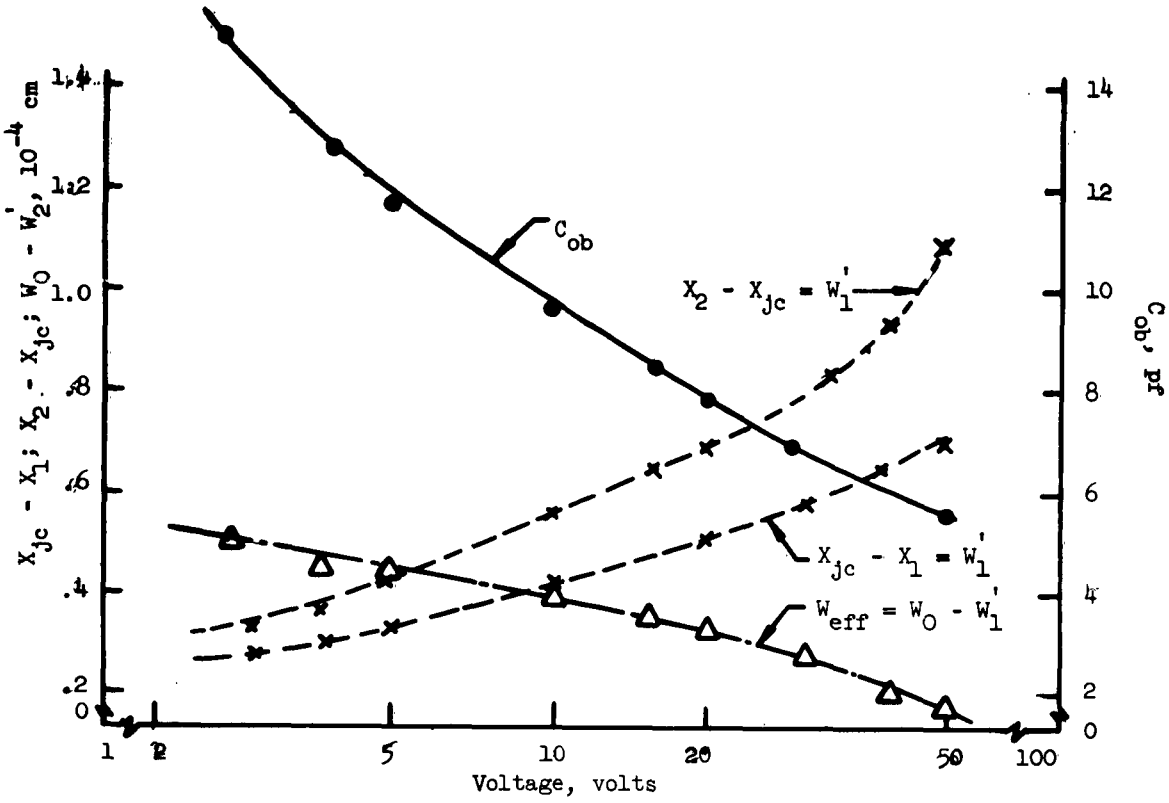


FIGURE 33 SPACE-CHARGE LAYER, EFFECTIVE BASE WIDTH AND COLLECTOR CAPACITANCE FOR GaAs PNP POWER TRANSISTOR

net impurity concentration in the base and with the application of a few more volts punchthrough will occur. At the same time, the maximum field has reached a value of about 4×10^5 volts/cm and at this point breakdown will occur.

Thus punchthrough and breakdown are optimized and occur at:

$$V_{PT} \approx BV_{CBO} \approx 50 \text{ volts}$$

(4) Transport Factor and Current Gain

Recombination losses in the bulk affect the transport factor and, therefore, the current gain. As is known, the minority carrier lifetime in gallium arsenide is rather limited ($\sim 10^{-9}$ sec) and hence the diffusion length can generally not be considered much larger than the base width.

The transport factor,

$$\beta_o = \frac{I_C}{I_{PO}} = \cosh - 1\ (W_{eff}/L_P) \qquad (18)$$

represents the ratio of the collector current to the injected hole current.

In particular, assuming a lifetime $\tau_p = 3 \times 10^{-9}$ sec and with $D_p = 9.6\ cm^2/sec$ we obtain a diffusion length,

$$L_P = \sqrt{D_P\ \tau_p} = \sqrt{9.6 \times 3 \times 10^{-9}} = 1.7 \times 10^{-4} cm$$

Thus, with $W_{eff} = 0.29 \times 10^{-4}$ cm at $V_C = 20$ volts

$$\beta_o = \cosh -1\ (0.17) = 0.988$$

Therefore, the current gain in a common base configuration is:

$$\alpha = \beta \gamma \delta = 0.972 \qquad (19)$$

if the previously calculated value of $\alpha = 0.987$ is used and no multiplication occurs ($\delta = 1$).

Finally, changing to a common emitter configuration, we obtain a current gain of:

$$h_{fe} = \frac{\alpha}{1 - \alpha} = \frac{0.972}{1 - 0.972} = 34.5$$

b. NPN-Approach

The design calculations for the npn-approach are similar to those for the pnp-version. The use of epitaxial material and a narrower base width is assumed.

(1) Impurity profile

The npn power transistor makes use of the impurity distribution shown in Figures 34 and 35. The epitaxial substrate, doped n-type to a level of 10^{18} carriers/ cm^3 or more, is not shown. The epitaxial layer itself is n-type doped to 2×10^{15} carriers/ cm^3 with a grown

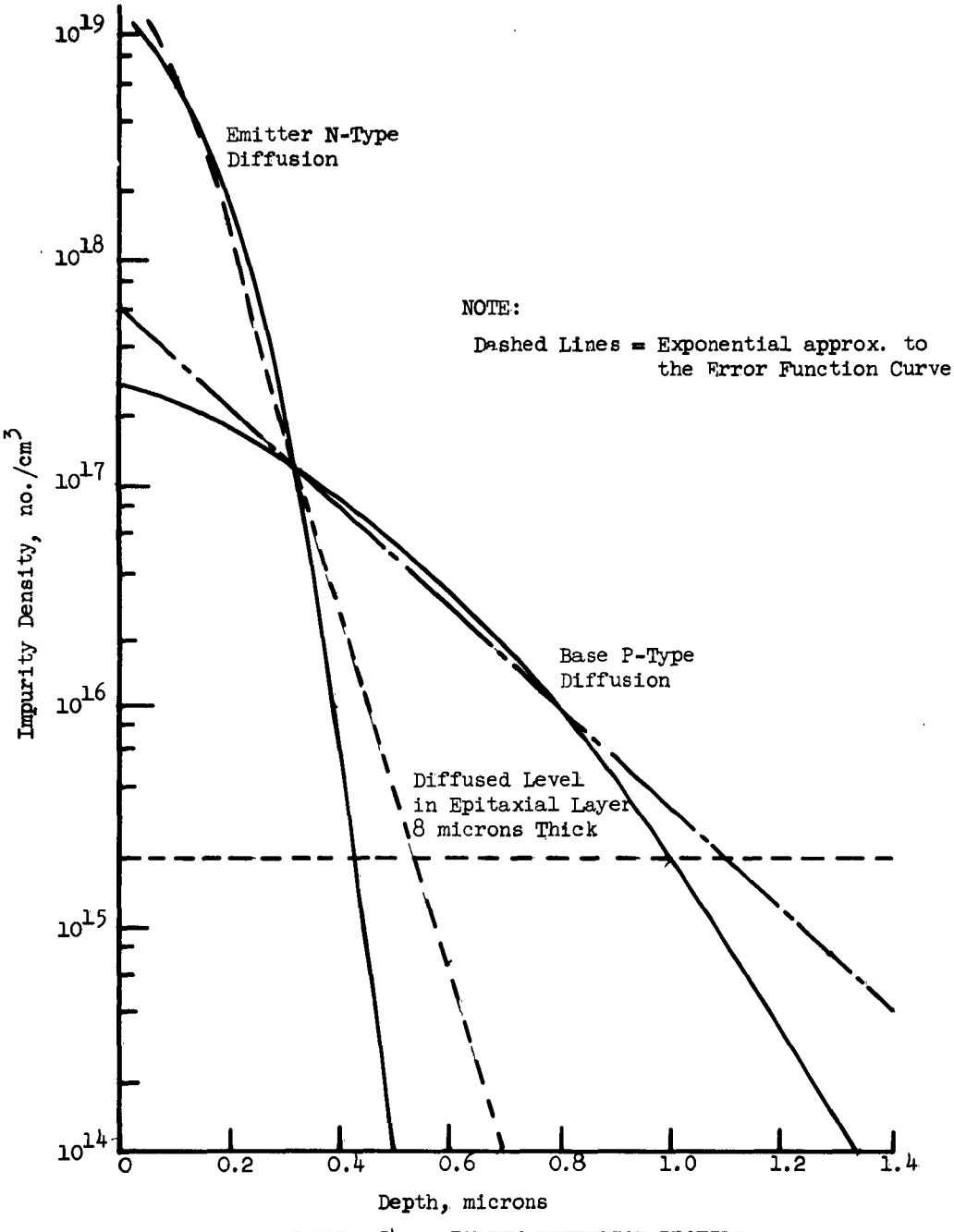


FIGURE 34 IMPURITY DIFFUSION PROFILE

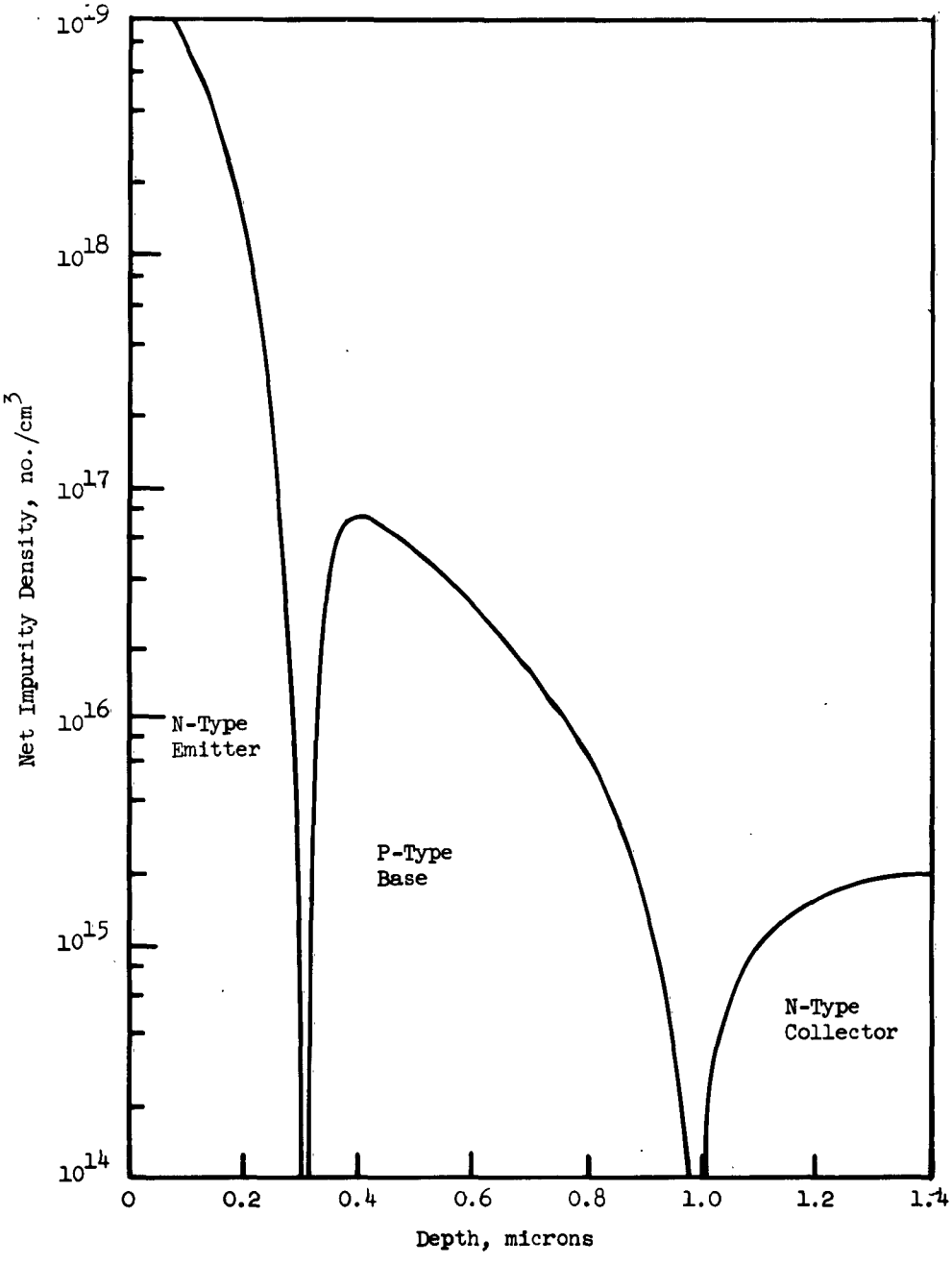


FIGURE 35 NET IMPURITY PROFILE

region approximately eight microns thick. The p-type base region has a surface concentration of 3×10^{17} carriers / cm^3 and a junction depth one micron below the surface. The n-type region is doped to a surface concentration of 2×10^{19} carriers/cm^3 for a junction depth of 0.3 microns. An error function distribution is assumed.

The quantities for equation (11) are:

$N_{10} = 2 \times 10^{19}$ carriers/ cm^3 = Tin surface concentration

$N_{20} = 3 \times 10^{17}$ carriers/ cm^3 = Zinc surface concentration

$N_3 = 2 \times 10^{15}$ carrier/ cm^3 = Epitaxial layer n-type doping level

$L_1 = 1.55 \times 10^{-5}$ cm = Diffusion length for tin emitter diffusion

$L_2 = 5.21 \times 10^{-5}$ cm = Diffusion length for Zinc base diffusion

We can evaluate:

$$J_1 = \int_{0.3\mu}^{1.0\mu} -N'(x)\,dx = 1.36 \times 10^{12}/cm^2$$

and

$$J_2 = \int_{+0.1\mu}^{0.3\mu} N'(x)\,dx = 2.93 \times 10^{13}/cm^2$$

$$\gamma = 1 + \frac{8}{100} \quad \frac{1.36 \times 10^{12}}{2.93 \times 10^{13}} = 0.996$$

which is quite satisfactory.

The base transport factor can be written as a first approximation

$$\beta = \frac{1}{1 + \frac{1}{2}(W_{eff}/Ln)^2} \qquad (20)$$

where W_{eff} is the effective base width with collector bias applied and L_n is the minority carrier diffusion length in the base region. For an operating collector bias of 30 volts, the effective base width

is found to be 0.23 microns. The diffusion length L_n can be calculated, for an expected lifetime of 10^{-10} seconds, as

$$L_n = \sqrt{D_n \tilde{\ }_n} = \sqrt{100 \times 10^{-10}} = 10^{-4}\text{cm}$$

so that

$$\beta = \frac{1}{1 + \frac{1}{2}\left(\frac{0.23}{1.0}\right)^2} = 0.974$$

which is also quite satisfactory.

It is next necessary to establish values for the punchthrough voltage and avalanche breakdown voltages at the collector and emitter. The calculation of punchthrough voltage is most easily accomplished using the design curves of Figure 32. Using the epitaxial layer doping of 2×10^{15} carriers/cm^3, the p-type impurity diffusion length of $L_2 = 5.21 \times 10^{-5}$cm and the normalized junction depth:

$$Y_{jc} = \frac{X_{jc}}{L_2} = \frac{10^{-4}}{5x21 \times 10^{-5}} = 1.92$$ we calculate the normalizing field and voltage

$$E_N = 4.42 \times 10^3 \text{ v/cm}$$

and

$$V_N = 0.06 \text{ volts}$$

We are seeking the punchthrough voltage where

$$V_{pt} = 2\, Y_{jc}\, (Y_{pt} - Y_{jc}) = 4.41$$

so that

$$V_{PT} = 3 \times 10^3\, V_N$$

and

V_{PT} = 180 volts

The extent of the collector depletion layer can also be established using Figure 32. A typical collector voltage of 30 volts will be used so that

$$\frac{V}{VN} = 5 \times 10^2$$

From Figure 32 $\mu_2 = 3.45$, and $\mu_1 = 29$. We can write

$$y = \frac{\mu + 2Y_{jc}}{2Y_{jc}}^2 = \frac{x}{L_2} \qquad (21)$$

so that the base edge of the collector depletion layer is found to be 0.532 microns from the surface so that the effective base width is 0.532-0.30 = 0.23 microns. The collector edge of the collector depletion layer is found to be 5.08 microns from the surface so that the collector depletion layer width is 5.08 - 0.53 = 4.55 microns.

The total collector capacitance can now be calculated

$$C_c = \frac{\epsilon A}{W} = \frac{0.98 \text{ pf/cm} \times 1.04 \times 10^{-3} \text{cm}^2}{4.55 \times 10^{-4} \text{ cm}} = 2.25 \text{ pf.}$$

Avalanche breakdown voltage in gallium arsenide has been established as a function of doping density over a limited doping region. At a doping level of 2×10^{15} carrier / cm^3 the breakdown voltage is greater than 100 volts for the collector junction. The breakdown voltage at a 10^{17} carrier / cm^3 doping level is greater than 8 volts for the emitter junction.

The basic figure of merit for high frequency transistors is given by the equation below

$$(PG)^{1/2} f = \frac{1}{4\pi} \sqrt{\frac{1}{r_{bb'} C_c \tau_{ec}}} \qquad (22)$$

where PG is the power gain of the device, f is the frequency of operation r_{bb}' is the base lead resistance, C_c is the collector capacitance and τ_{ec} is the total charge transfer time from emitter to collector.

The charge transfer time is made up of three components:

1) the emitter junction charging time, $\tau_e = r_e C_{Te}$,

2) the base transit time, $\tau_b \leq \frac{W_{eff}^2}{2Dn}$ and

3) the collector depletion layer transit time, $\tau_c = \frac{xm}{2V_{sat}}$

The value of the emitter incremental resistance r_e is given by

$$r_e \quad \frac{n\,k\,T}{q\,I_e} = \frac{2 \times 0.025\ v}{0.20\ amp} = 0.25\ ohms \qquad (23)$$

The emitter capacitance is

$$C_{Te} = \left(\frac{q \epsilon N}{2(\phi - V)} \right)^{1/2} \quad Ae = 83\ pf \qquad (24)$$

so that

$$\tau_e = r_e\, C_{te} = 0.25 \times 83 \times 10^{-12} = 20.7 \times 10^{-12}\ sec$$

The base transit time is given by

$$\tau_b \leq \frac{W_{eff}^2}{2\,D_n} \qquad \frac{(2.3 \times 10^{-5})^2}{2 \times 100} = 2.65 \times 10^{12}\ sec$$

The collector depletion layer transit time is given by the width of the depletion layer divided by twice the saturation velocity for carriers in the semiconductor. The saturation velocity is determined by lattice-scattering and no value is yet established for gallium arsenide. The known value for silicon, 5×10^6 cm/sec., will be used here so that

$$\tau = \frac{Xm}{2V_{sat}} = \frac{4.55 \times 10^{-4}}{2 \times 5 \times 10^6 cm/sec} = 45.5 \times 10^{-12}\ sec.$$

The total transfer to collector charge transfer time is

$$\tau_{ec} = (20.7 + 2.65 + 45.5) \times 10^{-12} \text{ sec } 69 \times 10^{-12} \text{ sec.}$$

It should be noted that the major contribution to this time constant comes from the collection depletion layer transit time and that this is dependent on a lattice-scattering limited velocity which is chosen arbitrarily equal to that for silicon.

The value of r_{bb}' is determined from the sheet resistance of the diffused base layer and the geometric structure. The value of the sheet resistance for the base diffusion (excluding the region under the emitter since the current flows only at the emitter edges) is obtained from Figure 35. A surface concentration of 3×10^{17} carriers / cm^3 with a substrate resistivity less than 10^{16} carriers / cm^3 gives a value of 3.8 for $1/\rho_s x_j$. Since the collector junction depth is 10^{-4}cm, the sheet resistance is equal to 2.63×10^3 ohms/square. This gives for the geometry of Figure 30.

$$r_{bb}' = \frac{\rho_\square}{2 \times 16} = 8.23\Omega$$

The Gain-Bandwidth product with the values found in (equation 22)

$$(PG)^{1/2} \times f = \frac{1}{4\pi} \left(\frac{1}{8.23 \times 2.25 \times 10^{-12} \times 69 \times 10^{-12}} \right)^{1/2} =$$

2.5×10^9 cps

which exceeds the power gain of 11 db required at 50 mc by far.

3. Device Processing

The power transistor starting material is single crystal gallium arsenide oriented to the (111) face with carrier concentration of between 1×10^{16} and 6×10^{16} carriers / cm^3. Although accurate orientation is not necessary

the crystal is customarily oriented within 0.5° of the (111) face.

Since the processing order and many of the techniques are the same for pnp and npn structures, they will be described together and differences noted when required. Figure 36 shows the processing sequence for the pnp structure and Figure 37 shows the sequence for the npn structure. One of the major differences is in the technique for diffusing the base region. For the pnp structure, the base is n-type and is formed by diffusing sulfur from a very limited source in a sealed quartz ampoule. For this diffusion, it is necessary to have about one-half atmosphere of arsenic pressure in the ampoule at the diffusion temperature (900-950°C). A typical diffusion at 900°C for one hour gives a junction depth of .04 to .05 mils and a surface concentration of about 8×10^{17} carriers/cm^3. For the npn structure, the only suitable p-type base diffusant was found to be zinc. Here, however, control of the surface concentration in the desired range of 2 to about 5×10^{17} carriers/cm^3 cannot be achieved by conventional diffusion methods. Therefore, a multiple diffusion process using a limited source has been developed, taking advantage of the facts that the diffusion constant for zinc is concentration dependent, and that silicon dioxide does not mask against zinc.

The base diffusion is carried out in three steps: 1) Starting with a silicon dioxide coated wafer, a very limited amount of zinc is deposited into the oxide layer. This deposition is performed at 750°C for four minutes using a .25% zinc in gallium source. This process may be called deposition. 2) The zinc is then diffused from the oxide into the skin of the gallium arsenide. The wafer is sealed in an evacuated quartz ampoule. A diffusion time of about 4 hours at 800°C will yield a sheet resistance in the neighborhood of 300 ohms/square. 3) Subsequently, the doped oxide is etched off

a) N-type base layer formed by sulfur diffusion

b) Deposition of diffusion mask (2 SiO_2 layers)

c) Definition of emitter pattern in SiO_2 by photoresist methods

d) P-type emitter diffusion (zinc)

e) Emitter contact evaporation (Ag) through metal mask, and sintering

f) Removal of SiO_2, evaporating base contact stripes (silver-tin) and sintering

g) Wax evaporation through metal mask and mesa etching

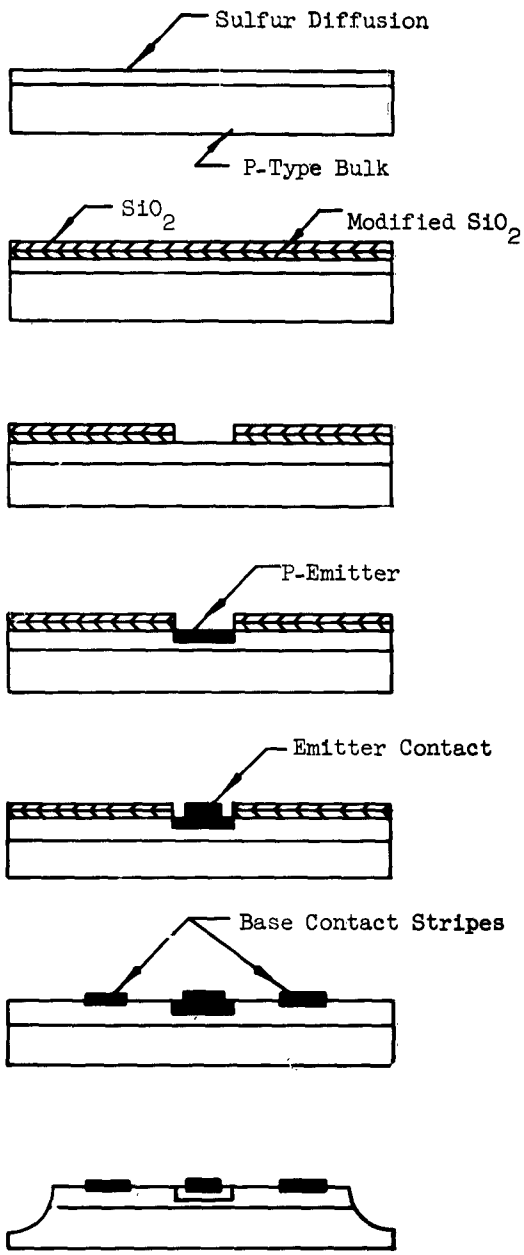


FIGURE 36 BASIC PROCESS STEPS FOR PNP-POWER TRANSISTOR

a) Zinc deposition in SiO_2 layer

b) Zinc diffusion

c) Definition of emitter pattern in SiO_2 by photoresist methods

d) N-type emitter diffusion (tin)

e) Emitter contact evaporation (Ag-Sn) through metal mask and sintering

f) Removal of SiO_2 from base contact area. Evaporation and sintering of base contact (silver)

g) Wax evaporation through metal mask and mesa etching

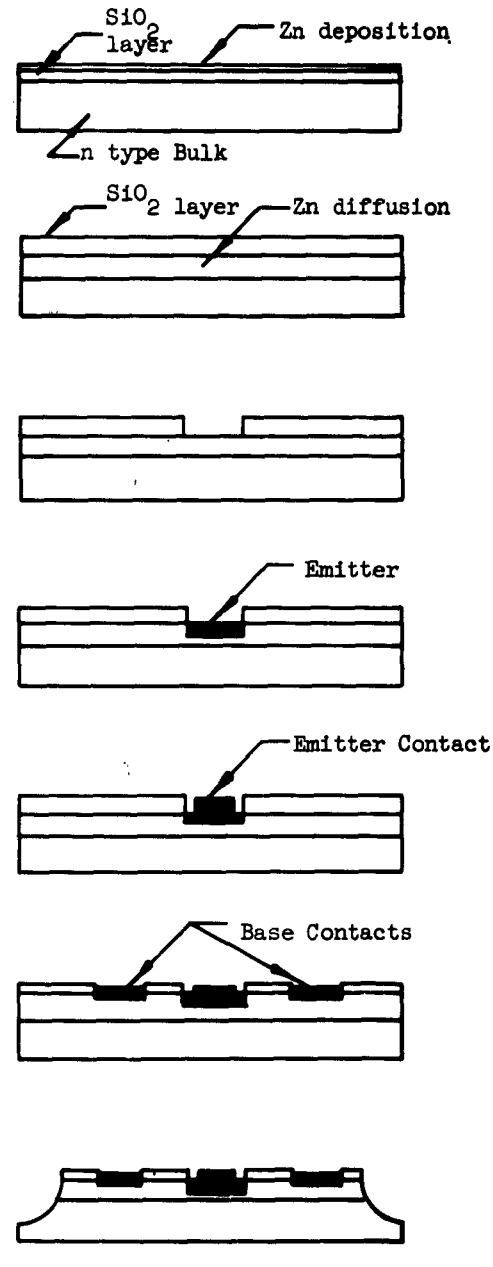


FIGURE 37 BASIC PROCESS STEPS FOR NPN-POWER TRANSISTOR

and replaced by a plain oxide layer. A redistribution step at 900°C for 24 hrs. yields the desired junction depth of 0.8×10^{-4} cm whereby the sheet resistance of 3000 ohms/square is maintained.

It should be pointed out that this technique offers a unique control of doping and is the only way known so far for processing a double diffused gallium arsenide transistor structure which shows current and power gain. Furthermore, the solid-to-solid diffusion (step 2) provides the possibility, of achieving a planar p-type junction, if only doped silicon dioxide blocks are left on the surface and a protective plain layer is put over the entire wafer.

After base formation, a silicon dioxide layer is applied to the pnp transistor wafer (the npn wafer already has an oxide coating from the zinc diffusion). A rectangular opening is etched into the oxide layer using standard photolithographic techniques. The oxide layer on the pnp structure is a double layer with the under layer containing a modified form of silicon dioxide. For the pnp structure, zinc is diffused at about 750°C. Diffusion time is typically a few minutes and is adjusted for each slice depending on the base diffusion depth. This diffusion is done in a sealed quartz ampoule using elemental zinc as the diffusant. For the npn structure, tin is diffused at a temperature of 900-950°C for a few minutes with the time adjusted once again for the base diffusion depth on each individual wafer. The surface concentration of the zinc emitter is about 10^{20} carriers per cm^3 and the surface concentration of the tin emitter is about 10^{19} to 2×10^{19} carriers per cm^3. As with all n-type diffusants, it is necessary to use about 1/2 atmosphere of arsenic in the sealed quartz diffusion ampoule while diffusing tin. The tin source is an evaporated metallic layer on the back of the wafer.

After emitter diffusion, the back of the wafer is lapped and a collector

contact is applied. The n-type contact is made by nickel plating over an evaporated tin layer and then sintering the coating in hydrogen at 600°C. The p-type contact is made by evaporating gold-1% manganese over a nickel plating and also firing in hydrogen at 600°C.

Emitter and base contacts are next applied to the transistor by vacuum evaporation of silver for the p-type contact and tin doped silver for the n-type contact. The silicon dioxide mask covering the wafer is completely removed from the pnp structure before applying emitter and base contacts, while on the npn structure, the silicon dioxide is removed only from the immediate contact area and thus remains on the surface everywhere except on the contact area. This oxide removal is done by the same technique as was used to remove it before the emitter diffusion step. After evaporation, the metal contacts are sintered at 600°C in hydrogen.

Mesa formation is a standard technique and involves evaporation of a wax pattern through a metal mask followed by etching to remove gallium arsenide everywhere except under the rectangular wax areas.

Scribing, mounting and bonding are done exactly as was described for the gallium arsenide switching transistor except that gold wire is used for all connections.

Etching and surface treatments of various kinds have been investigated but have not been used because of the high series resistance between emitter and base which occurs on etching. These units form an electrolytic cell and will etch rapidly on just being dipped into sodium or potassium hydroxide solutions with no external bias applied.

4. Measurements and Device Performance

Testing was performed on standard test equipment using procedures and

techniques worked out for similar silicon power transistors.

The collector and the emitter exhibit excellent diode characteristics. Figure 38 shows the collector reverse current for a typical unit. Figure 39 shows the collector forward current. Figure 40 and 41 show the emitter reverse and emitter forward current respectively. It may be noted that the emitter characteristic is considerably better than for an alloyed type gallium arsenide device.

The current gain on approximately 80 units tested was in the range from 1.5 to 7 for about 50% of the units with the center of the distribution at a beta value of 2 for I_C = 100 μa and V_C = 5 volts. One unit showed a maximum DC-current gain of 26. The collector family for this unit is given in Figure 42.

In Figure 43 the dependence of beta upon the collector current is shown. As can be seen from this graph the gain is higher than 10. in the range 0.1 to 100 ma. Complete data on 10 sample units are given in Table XVIII.

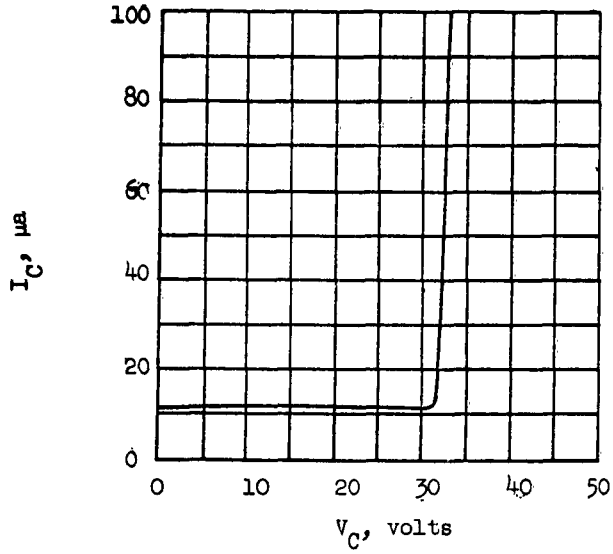


FIGURE 38 COLLECTOR REVERSE CHARACTERISTICS FOR DOUBLE-DIFFUSED NPN POWER TRANSISTOR

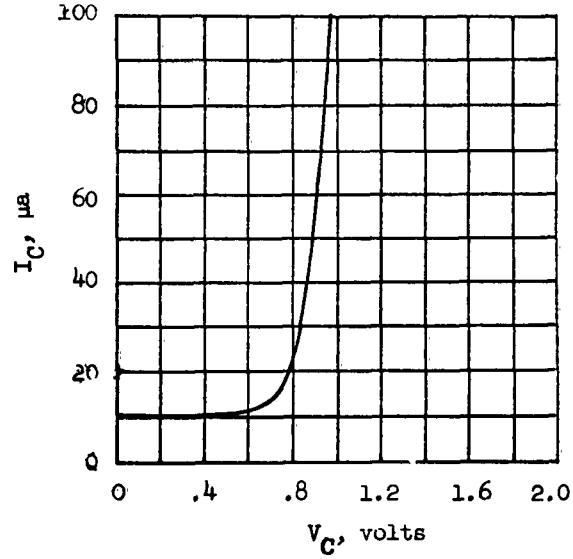


FIGURE 39 COLLECTOR FORWARD CHARACTERISTICS FOR DOUBLE-DIFFUSED NPN POWER TRANSISTOR

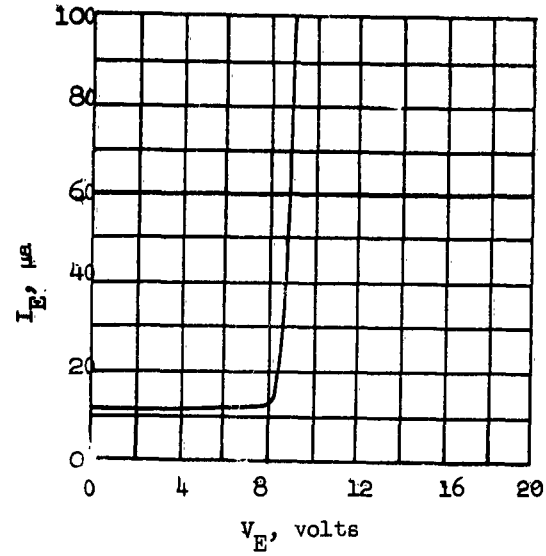


FIGURE 40 EMITTER REVERSE CHARACTERISTICS FOR DOUBLE-DIFFUSED NPN POWER TRANSISTOR

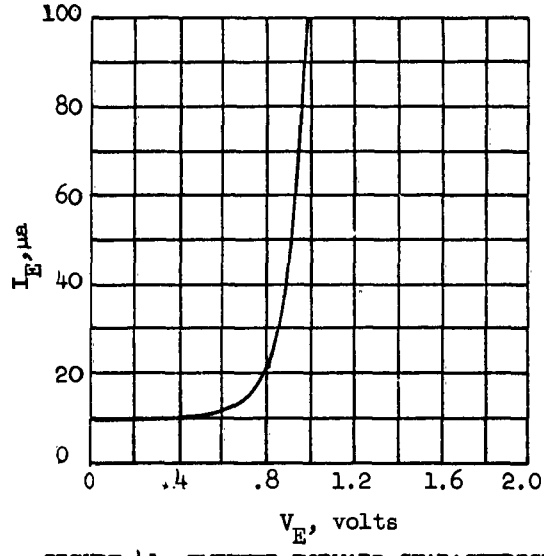
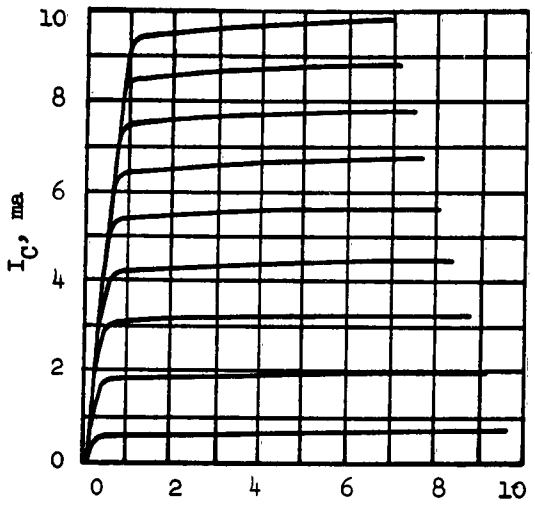


FIGURE 41 EMITTER FORWARD CHARACTERISTICS FOR DOUBLE-DIFFUSED NPN POWER TRANSISTOR



FIGRE 42 VOLTAGE-CURRENT CHARACTERISTIC OF HIGH CURRENT GAIN NPN UNIT

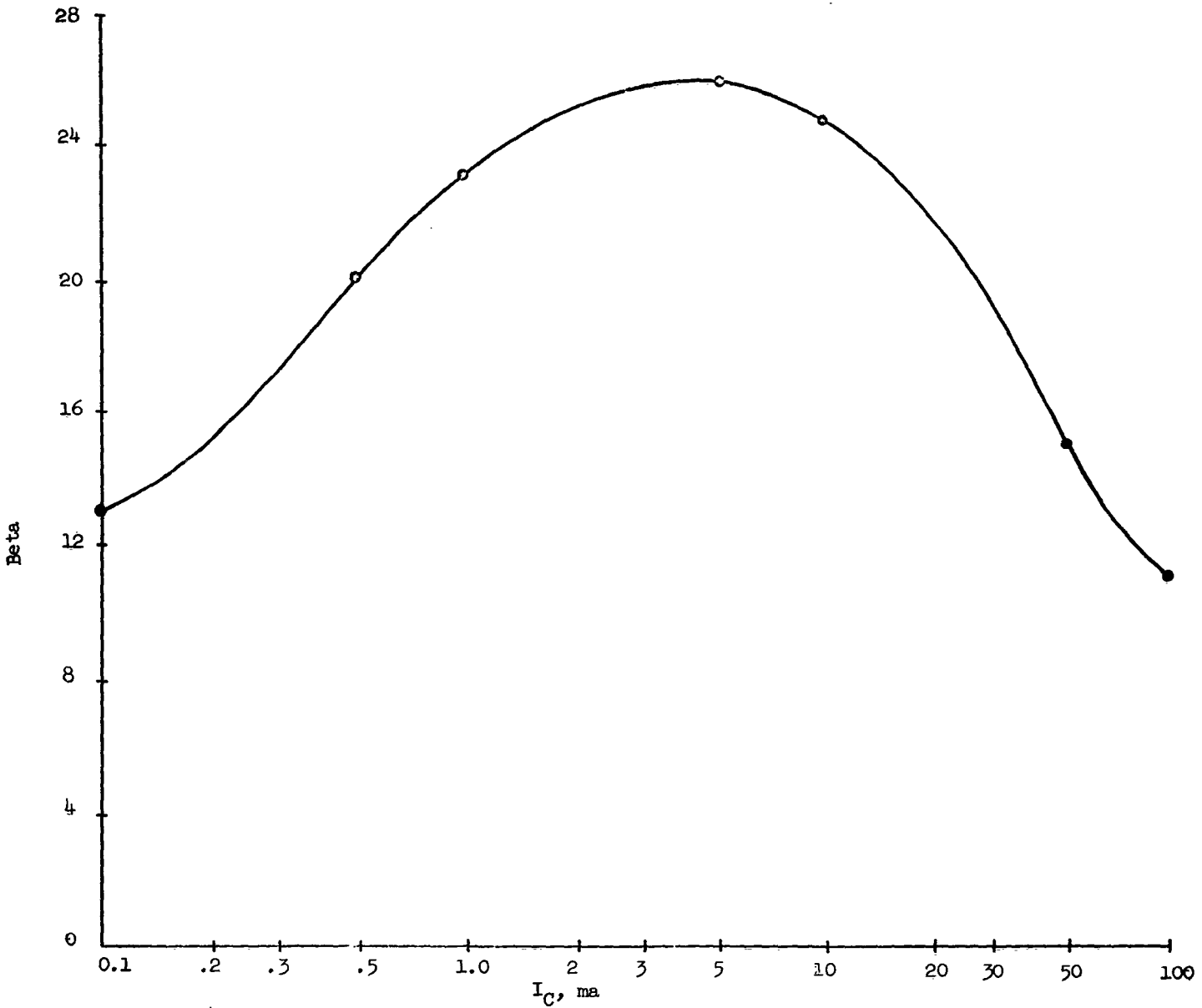


FIGURE 43 BETA VERSUS COLLECTOR CURRENT FOR HIGH CURRENT GAIN NPN UNIT

TABLE XVIII

CHARACTERISTICS OF NPN POWER TRANSISTOR SAMPLES

Unit No.	BV_{CBO} I_C=1ma, volts	BV_{EBO} I_E=100μa, volts	h_{FE} I_C=50ma	100 mcs h_{fe} I_C = 50ma	C_{ob} V_{CB}=10v, pf	Power Gain, db f = 50 mc V_{CE}=15v I_C=30ma T=25°C	Power Gain, db f = 50 mc V_{CE}=15v I_C=30ma T=200°C	Power Gain, db f = 50 mc V_{CE}=12v I_C=60ma T=25°C
1	43	9.8	2.5	1.6	9.3	6.5	12.0	9.2
2	22	9.0	2.2	1.7	7.0	7.2	15.2	10.6
3	52	5.5	2.5	1.7	7.8	6.0	9.5	8.5
4	51	9.5	3.0	1.2	7.3	2.8	11.2	5.7
5	47	9.4	2.5	1.3	6.2	5.0	14.2	8.2
6	23	7.5	3.2	1.25	7.6	3.2	8.9	5.8
7	28	10.2	3.0	1.5	6.0	5.6	11.8	7.4
8	24	10.2	3.0	1.9	6.4	4.8	10.4	6.1
9	23	10.5	3.5	2.3	7.0	4.5	11.4	7.2
10	23	10.8	3.7	2.3	7.6	6.1	14.2	8.2

L. Future Uses for Gallium Arsenide

One of the most attractive future applications for gallium arsenide is in devices for the generation of infrared light both as a simple light source and as an injection laser. It has been known for some time that forward biased pn junctions in gallium arsenide emit infrared light with a high efficiency. Considerable study has gone into a determination of the spectrum of this light, the mechanism for its generation and the possible applications of such a light source. At low temperatures, a very intense radiation is emitted at about 8600 angstroms. This wavelength corresponds to an energy of about 1.47 ev., just a little less than the band gap energy for gallium arsenide. The light source can be made arbitrarily large or small by making the emitting pn junction large or small.

The presence of this intense infrared radiation in gallium arsenide related to injection current across a pn junction makes possible an injection laser. In a normal crystalline or gas laser, pumping is accomplished by illuminating the laser with an intense burst of broad band light. A very small percentage of the energy is returned as a coherent, monochromatic beam of light, the wavelength of which is determined by the properties of the laser material. Only very recently, it has become possible to achieve laser action in a gallium arsenide crystal by very carefully polishing two plane parallel surfaces perpendicular to the pn junction and injecting a short high current pulse at the liquid nitrogen temperature (77°K).

Since the gallium arsenide injection laser is very new, a number of improvements should come from further development work. Continuous laser action at much lower current densities may be possible. One of the major advantages from a continous injection laser is the ease of modulation of the laser beam by electrically modulating the input power to the pn junction.

Gallium arsenide pn junction also have a very fast recovery time due to the

extremely small minority carrier lifetime. This characteristic makes photo diodes possible with speeds well into the kmc range. Therefore, gallium arsenide injection laser and gallium arsenide photo diodes are extremely compatible and complementary.

Early samples of the gallium arsenide low power rectifiers were supplied to manufacturers of jet engines for use in the very high ambient temperature encountered near these engines. Silicon rectifiers were used for this application, but could not be mounted in the desired location, close to the engine. The silicon rectifiers have to be mounted some distance from the engine in a cooling stream of fuel. This positioning is both complicated and undesirable. The earliest gallium arsenide samples operated satisfactorily in the high ambient temperatures close to the engine. However, difficulty arose due to cracked pellets which were encounted after many temperature cycles as the engines were heated and cooled. As a result of this problem the rectifiers were modified by changing both the high temperature solder and the processing used to solder the pellets to the enclosure. This change greatly improved the rectifier in several ways--reduction of cracking, better mechanical strength and lower thermal resistance. Samples of the improved units operated without difficulty, close to the jet engine. The temperature to which these rectifiers are exposed is in excess of 300°C. Samples of these improved rectifiers also passed many accelerated tests designed to prove the reliability of the device.

Recent data has proved that gallium arsenide solar cells have a greater resistance to radiation than either p/n or n/p silicon solar cells. For 17.5 mev and 95 mev protons, gallium arsenide solar cells are superior in radiation resistance to n/p silicon cells by a factor of ten. N/p silicon cells in turn are superior to p/n silicon cells. This means that for the same initial conversion efficiency gallium arsenide cells will take ten times as much radiation flux as n/p silicon cells before falling to some predetermined value.

(In most of the tests the critical flux is defined as that to reduce the maximum power output by 25 percent). For electron energies of about 800 ev, the critical flux for both types of cells is about the same; but for electron energies in the neighborhood of 5.5 mev, the ten to one advantage for gallium arsenide is again obtained.

Gallium arsenide solar cells are now being tested on several satellites. Solar cells on the Anna B Satellite have shown very little efficiency change over a period of several months. These cells have a cover glass about .006-inch thick. The change in efficiency of these cells compares with the change in a n on p silicon cell with a cover glass .060-inch thick. The p on n silicon solar cells deterioriated almost an order of magnitude more than either the gallium arsenide or the n on p silicon cell. Another group of gallium arsenide solar cells were used in the first Relay Satellite. These cells had no cover glass and as a result were damaged by low velocity electrons and protons. Gallium arsenide solar cells with a cover glass will be on the second Relay Satellite and other satellites soon to be launched.

II. CONCLUSIONS

Analysis indicated that gallium arsenide is the most promising semiconductor material in terms of high-frequency, high-temperature performance. It was found that obtainable operating temperatures with gallium arsenide were higher than those with silicon and that these temperatures could be reached without sacrificing the high-frequency performance of germanium. It was also found that gallium arsenide ranks very high as a material to be used for the conversion of solar energy by photovoltaic devices, and will probably give the highest realizable conversion efficiency. In addition, since the quantum efficiency of forward current in gallium arsenide p-n junctions is high, this material would make unique lasers, capable of operating electrically by the injection of minority carriers rather than by optical pumping.

The design used for the low-, medium-, and high-power rectifiers was capable of producing units that met the objective specifications set forth by this contract. These units were also life tested at 400°C and yielded extremely good results.

Since the operating voltage for gallium arsenide Zener diodes and tunnel diodes are higher than those in similar silicon or germanium devices, they are capable of fulfilling some applications for which there was no available semiconductor device.

The results obtained with unipolar transistors were not especially satisfactory. This is because the device technology at the time these units were produced was rather crude. A similar attempt at producing these transistors today would yield results far superior to those previously obtained.

Switching and power transistors were also produced under this contract. The devices are capable of meeting all the objective specifications of the

contract except for current gain. These devices are better than either germanium or silicon with respect to high-frequency at high-temperature operation.

The most significant accomplishment of this contract was the development of gallium arsenide tunnel diodes and varactor diodes. These diodes have many advantages over similar germanium and silicon devices and have definitely advanced the state-of-the-art in these areas.

Aeronautical Systems Division, Dir/Avionics, Electronic Technology Lab, Wright-Patterson AFB, Ohio.
Rpt No. ASD-TDR-63-59, DEVELOPMENT OF HIGH TEMPERATURE SEMICONDUCTOR DEVICES. Final report, Feb 63, 110p. incl illus., tables, 12 refs.
Unclassified Report

Gallium arsenide was chosen as the semiconductor material best suited to meet the needs of high-temperature operation because of its wide band gap and high electron mobility. Even though gallium arsenide has the best potential for high-temperature operation, it does have a few limitations. One of these being that there are a number of major differences between elemental and compound semi-

(over)

1. Semiconductor devices
2. Semiconductor materials
3. Diodes
4. Rectifiers
5. Transistors
I. AFSC Project 4460, Task 446002
II. Contract AF33(600)-37267
III. Radio Corporation of America, Semiconductor and Materials Division, Somerville, N. J.
IV. H. Becke, et al.
V. Aval fr OTS
VI. In ASTIA collection

Aeronautical Systems Division, Dir/Avionics, Electronic Technology Lab, Wright-Patterson AFB, Ohio.
Rpt No. ASD-TDR-63-59, DEVELOPMENT OF HIGH TEMPERATURE SEMICONDUCTOR DEVICES. Final report, Feb 63, 110p. incl illus., tables, 12 refs.
Unclassified Report

Gallium arsenide was chosen as the semiconductor material best suited to meet the needs of high-temperature operation because of its wide band gap and high electron mobility. Even though gallium arsenide has the best potential for high-temperature operation, it does have a few limitations. One of these being that there are a number of major differences between elemental and compound semi-

(over)

1. Semiconductor devices
2. Semiconductor materials
3. Diodes
4. Rectifiers
5. Transistors
I. AFSC Project 4460, Task 446002
II. Contract AF33(600)-37267
III. Radio Corporation of America, Semiconductor and Materials Division, Somerville, N. J.
IV. H. Becke, et al.
V. Aval fr OTS
VI. In ASTIA collection

conductors.

A number of different devices were developed under this contract. These include: switching diodes, low-power rectifiers, medium-power rectifiers, high-power rectifiers, zener diodes, tunnel diodes, unipolar transistors, switching transistors, power transistors and solid ceramic circuits. Of these the most successful were tunnel diodes and varactor diodes, since they are able to fulfill the requirements of some specific applications not filled by similar germanium and silicon devices.

conductors.

A number of different devices were developed under this contract. These include: switching diodes, low-power rectifiers, medium-power rectifiers, high-power rectifiers, zener diodes, tunnel diodes, unipolar transistors, switching transistors, power transistors and solid ceramic circuits. Of these the most successful were tunnel diodes and varactor diodes, since they are able to fulfill the requirements of some specific applications not filled by similar germanium and silicon devices.